

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1359	100.0	258	21	AAV94893	Human protein clon
2	1359	100.0	272	21	AAV57113	Human prostate can
3	1326	97.6	296	20	AAV13380	Amino acid sequenc
4	1326	97.6	296	22	AAU29039	Human PRO polypept
5	1326	97.6	296	22	AAV80248	Human PRO270 prote
6	1326	97.6	329	22	ABV12004	Human CGI-31 prote
7	1322	97.3	296	19	AAW75204	Human secreted pro
8	1313	96.6	293	19	AAW75226	Human secreted pro
9	1308	96.2	296	23	ABV06300	Prostaglandin E2 (
10	1259	92.6	372	22	AAV88335	Human membrane or

PR 09-SEP-1998; 98JP-0254736.
PR 29-SEP-1998; 98JP-0275505.
XX (SAGA) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX Kato S, Kimura T;
XX WPI; 2000-182694/16.
DR Novel human proteins having hydrophobic domains useful for treating
XX osteoporosis, Alzheimer's disease, Parkinson's disease, asthma,
PT multiple sclerosis, rheumatoid arthritis, cancer, anaemia, and stroke -
PT Claim 1; Page 319-320; 351pp; English.
PS This sequence represents a human protein of the invention, which has
XX hydrophobic domains. The DNA sequences can be used as a probe or as a
CC genetic marker. The protein can also be used as a marker, and to identify
CC potential genetic disorders. The DNA and protein can also be used as
CC nutritional sources or supplements. The protein exhibits cytokine, cell
CC proliferation, cell differentiation activities and induces production of
CC other cytokines in certain cell populations. The protein also exhibits
CC immunostimulating or immune suppressing activity. It can be used in the
CC treatment of various immune deficiencies and disorders, and to treat
CC infectious diseases caused by viral, bacterial, fungal or other
CC infections. The protein is also used for treating autoimmune disorders
CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid
CC arthritis. It is also useful in the treatment of allergic reactions and
CC conditions such as asthma, and in immune suppression after organ
CC transplantation. The protein is useful in regulation of haematopoiesis
CC and consequently in the treatment of myeloid or lymphoid cell
CC deficiencies. It is also used in compositions for tissue growth or
CC regeneration. The protein is also used in the treatment of osteoporosis
CC or osteoarthritis and in the treatment of periodontal disease and other
CC tooth repair processes. The protein is used in the treatment of nervous
CC system disorders such as Alzheimer's disease, Parkinson's disease, and
CC Huntington's disease. They are useful for protection or regeneration and
CC treatment of lung or liver fibrosis, reperfusion injury in various
CC tissues, and conditions resulting from systemic cytokine damage. They are
CC also used for promoting or inhibiting tissue differentiation. They are
CC also used as contraceptives since they exhibit activin or inhibin related
CC activities and as a fertility inducing therapeutic. They are used for
CC treating various coagulation disorders and in treatment and prevention of
CC conditions resulting from coagulation activities e.g. myocardial
CC infarction or stroke. They also acts as receptors, receptor ligands or
CC inhibitors or agonists of receptor/ligand interactions. They are used to
CC treat inflammatory conditions such as septic shock, sepsis, ischaemia
CC reperfusion injury, arthritis, and nephritis. They can be used to
XX prevent tumours.

Query Match 100.0%; Score 1359; DB 21; Length 258;
Best Local Similarity 100.0%; Pred. No. 5.5e-140;
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVLAPLIALVYVPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNPCD 60
DB 1 MAVLAPLIALVYVPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNPCD 60
QY 61 FDWREVEILMFLSAIVMKNRRSMFLMTCKPPLYMGPEYIKYFNKTDDEELERKRVTW 120
DB 61 FDWREVEILMFLSAIVMKNRRSMFLMTCKPPLYMGPEYIKYFNKTDDEELERKRVTW 120
QY 121 IVEFANNSDCQSFAPYIADLSKYNCTGLNFGKVDGCRYTDVSTRYKVSPTSPTKQLP 180
DB 121 IVEFANNSDCQSFAPYIADLSKYNCTGLNFGKVDGCRYTDVSTRYKVSPTSPTKQLP 180
QY 181 TLILFQGGKEAMRRPQIDKGRAVSWTFSEENVIREFNINELYQRAKLSKAGDNIPREQ 240
DB 181 TLILFQGGKEAMRRPQIDKGRAVSWTFSEENVIREFNINELYQRAKLSKAGDNIPREQ 240

QY 241 PVASTPTTVSDGENKKDK 258
DB 241 PVASTPTTVSDGENKKDK 258
RESULT 2
AAB57113
ID AAB57113 standard; Protein; 272 AA.
XX AAB57113;
XX 13-MAR-2001 (first entry)
XX Human prostate cancer antigen protein sequence SEQ ID NO:1691.
XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease.
XX Homo sapiens.
OS WO200055174-A1.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000WO-US05988.
XX 12-MAR-1999; 99US-0124270.
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX Rosen CA, Ruben SM;
XX WPI; 2000-587513/55.
DR N-PSDB; AAF16316.
XX Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX Claim 11; Page 2161-2162; 2338pp; English.
XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX Sequence 272 AA;
XX Query Match 100.0%; Score 1359; DB 21; Length 272;
Best Local Similarity 100.0%; Pred. No. 5.9e-140;
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVLAPLIALVYVPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNPCD 60
DB 15 MAVLAPLIALVYVPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNPCD 74
QY 61 FDWREVEILMFLSAIVMKNRRSMFLMTCKPPLYMGPEYIKYFNKTDDEELERKRVTW 120

Db 75 FDWEVEILMFLSAIVMMKNRRSMFLMTCKPPLYMGPEIKYFNDKTIDEELDRKRVTW 134
Qy 121 IVEFFANNSDCOSFAPYIADLSLKYNCCTGLNFGKVDVGRYTDVSTRYKVSSTPLTKQLP 180
Db 135 IVEFFANNSDCOSFAPYIADLSLKYNCCTGLNFGKVDVGRYTDVSTRYKVSSTPLTKQLP 194
Qy 181 TLILFQGGKEAMRRQIDKKGRAVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQ 240
Db 195 TLILFQGGKEAMRRQIDKKGRAVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQ 254
Qy 241 PVASTPTTVSDGENKKDK 258
Db 255 PVASTPTTVSDGENKKDK 272
RESULT 3
Y13380
AAV13380 standard; Protein; 296 AA.
AC AAV13380;
XX
DT 25-JUN-1999 (first entry)
XX
DE
XX
XX Amino acid sequence of protein PRO270.
KW Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
KW anti-thrombotic; wound healing; tissue repair.
XX
OS Homo sapiens.
XX
XX
XX WO9914328-A2.
XX
PD 25-MAR-1999.
XX
PF 16-SEP-1998; 98WO-US19330.
XX
PR 25-NOV-1997; 97US-0066840.
PR 17-SEP-1997; 97US-0059113.
PR 17-SEP-1997; 97US-0059115.
PR 17-SEP-1997; 97US-0059117.
PR 17-SEP-1997; 97US-0059119.
PR 17-SEP-1997; 97US-0059121.
PR 17-SEP-1997; 97US-0059122.
PR 17-SEP-1997; 97US-0059184.
PR 18-SEP-1997; 97US-0059263.
PR 18-SEP-1997; 97US-0059266.
PR 15-OCT-1997; 97US-0062125.
PR 17-OCT-1997; 97US-0062285.
PR 17-OCT-1997; 97US-0062287.
PR 21-OCT-1997; 97US-0063486.
PR 24-OCT-1997; 97US-0062814.
PR 24-OCT-1997; 97US-0062816.
PR 24-OCT-1997; 97US-0063045.
PR 24-OCT-1997; 97US-0063120.
PR 24-OCT-1997; 97US-0063121.
PR 24-OCT-1997; 97US-0063127.
PR 24-OCT-1997; 97US-0063128.
PR 27-OCT-1997; 97US-0063329.
PR 27-OCT-1997; 97US-0063327.
PR 27-OCT-1997; 97US-0063541.
PR 28-OCT-1997; 97US-0063542.
PR 28-OCT-1997; 97US-0063544.
PR 28-OCT-1997; 97US-0063549.
PR 28-OCT-1997; 97US-0063550.
PR 28-OCT-1997; 97US-0063564.
PR 29-OCT-1997; 97US-0063435.
PR 29-OCT-1997; 97US-0063704.
PR 29-OCT-1997; 97US-0063732.

PR 29-OCT-1997; 97US-0063738.
PR 29-OCT-1997; 97US-0063734.
PR 29-OCT-1997; 97US-0064215.
PR 29-OCT-1997; 97US-0063735.
PR 31-OCT-1997; 97US-0063870.
PR 31-OCT-1997; 97US-0064103.
PR 03-NOV-1997; 97US-0064248.
PR 07-NOV-1997; 97US-0064809.
PR 12-NOV-1997; 97US-0065186.
PR 17-NOV-1997; 97US-0065846.
PR 18-NOV-1997; 97US-0065693.
PR 21-NOV-1997; 97US-0066120.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066772.
PR 24-NOV-1997; 97US-0066466.
PR 24-NOV-1997; 97US-0066770.
PR 24-NOV-1997; 97US-0066511.
PR 24-NOV-1997; 97US-00666453.
XX
XX (GETH) GENENTECH INC.
XX
XX Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
PI WPI; 1999-229533/19.
XX N-PSDB; AAX52251.
XX
XX New isolated human genes and polypeptides used in, e.g. treatment of
PT gastrointestinal ulceration
XX
XX Claim 12; Fig 76; 320pp; English.
XX
XX AAY13344-403 represent secreted and transmembrane human proteins.
CC The cDNA sequences are obtained from cDNA libraries, prepared from
CC fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.
CC The encoded polypeptides have specific uses based on their homology to
CC known polypeptides, e.g. PRO211 and PRO217 can be used for disorders
CC associated with the preservation and maintenance of gastrointestinal
CC mucosa and the repair of acute and chronic mucosal lesions
CC (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal
CC ulceration and congenital microvillus atrophy), skin diseases associated
CC with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial
CC cancers such as lung squamous cell carcinoma of the vulva and gliomas),
CC potent effects on cell growth and development, diseases related to
CC growth or survival of nerve cells including Parkinson's disease,
CC Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as
CC as a target for anti-tumor drugs. PRO533 may be used in the treatment
CC of Usher Syndrome or Atrophia areata; PRO269 can be used as an
CC anti-thrombotic agent; PRO287 polypeptides and portions may have
CC therapeutic applications in wound healing and tissue repair; PRO317 can
CC be used for treating problems of the kidney, uterus, endometrium, blood
CC vessels, or related tissue, e.g. in the heart of genital tract.
XX
SQ Sequence 296 AA;
Query Match 97.68; Score 1326; DB 20; Length 296;
Best Local Similarity 86.8%; Pred. No. 2.7e-136;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;
Qy 1 MAVLAPLIALVYSPRLSRWLQAPYLLSALLSAAFLLVKRLPPLCHGLPTQREDGNPCD 60
Db 1 MAVLAPLIALVYSPRLSRWLQAPYLLSALLSAAFLLVKRLPPLCHGLPTQREDGNPCD 60
Qy 61 FDWEVEILMFLSAIVMMKNRRS----- 83
Db 61 FDWEVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
Qy 84 -MFLMCKPPLYMGPEIKYFNDKTIDEELDRKRVTWIIEFFANNSDCOSFAPYIADL 142
Db 121 IVFLMCKPPLYMGPEIKYFNDKTIDEELDRKRVTWIIEFFANNSDCOSFAPYIADL 180
Qy 143 SLKYNCCTGLNFGKVDVGRYTDVSTRYKVSSTPLTKQLP TLILFQGGKEAMRRQIDKKGR 202

```
Db 181 SLKYNTGLNFGKVDVGRYDVSTRYKYSTPLTKOLPTLILFOGKEAMRRPQIDKGR 240
QY 203 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQVASTPTTVSDGENKKDK 258
    |||||
Db 241 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQVASTPTTVSDGENKKDK 296

RESULT 4
AAU29039
ID AAU29039 standard; Protein; 296 AA.
AC AAU29039;
XX
XX
DT 18-DEC-2001 (first entry)
DE Human PRO polypeptide sequence #16.
DE
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
XX Homo sapiens.
    WO200168848-A2.
PD 20-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US06520.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186968P.
PR 14-MAR-2000; 2000US-189320P.
PR 14-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 21-MAR-2000; 2000US-191314P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194647P.
PR 11-APR-2000; 2000US-195975P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196690P.
PR 11-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX
XX (GETH ) GENENTECH INC.
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
```

```
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2001-602746/68.
DR N-PSDB; AAS45940.
XX
PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX
XX Claim 11; Fig 32; 774pp; English.
XX
XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
XX The PRO polypeptides and their associated nucleic acids can be used to
XX detect the presence of a tumour in a mammal by comparing the level of
XX expression of a PRO polypeptide in a test sample of cells from the animal
XX and a control sample of normal cells, whereby a higher level of
XX expression in the test sample indicates the presence of a tumour in the
XX mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
XX and rabbits but are preferably human. The polypeptides can be used to
XX stimulate tumour necrosis factor (TNF) alpha release from human blood,
XX when contacted with it. A specific polypeptide can be used to stimulate
XX the proliferation or differentiation of chondrocyte cells. The PRO
XX proteins can be used to determine the presence of tumours and also
XX susceptibility to tumour development, particularly adrenal, lung, colon,
XX breast, prostate, rectal, cervical, or liver tumours, in mammalian
XX subjects. The oligonucleotide probes specific for the PRO nucleic acids
XX can be used for genetic analysis of individuals with genetic disorders.
XX
XX Sequence 296 AA;
    Query Match 97.6%; Score 1326; DB 22; Length 296;
    Best Local Similarity 86.8%; Pred. NO. 2.7e-136;
    Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;
QY 1 MAVLAPLIALVYSVPRLSRWLAQPYLLSALLSAAFLVRLKPLPLCHGLPTQREDGNPCD 60
    |||||
Db 1 MAVLAPLIALVYSVPRLSRWLAQPYLLSALLSAAFLVRLKPLPLCHGLPTQREDGNPCD 60
    |||||
QY 61 FDMREVEILMFLSAIVMKNNRS----- 83
    |||||
Db 61 FDMREVEILMFLSAIVMKNNRSITVEQHGIFNFMFSKVANTILFFRLDIRMGLLYITLC 120
    |||||
QY 84 -MFLMTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANNKNDQCSFAPYADL 142
    :|||||
Db 121 IVFLMTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANNKNDQCSFAPYADL 180
    |||||
QY 143 SLKYNTGLNFGKVDVGRYDVSTRYKYSTPLTKOLPTLILFOGKEAMRRPQIDKGR 202
    |||||
Db 181 SLKYNTGLNFGKVDVGRYDVSTRYKYSTPLTKOLPTLILFOGKEAMRRPQIDKGR 240
    |||||
QY 203 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQVASTPTTVSDGENKKDK 258
    |||||
Db 241 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQVASTPTTVSDGENKKDK 296
    |||||

RESULT 5
AAB80248
ID AAB80248 standard; Protein; 296 AA.
XX
XX AAB80248;
XX
XX 24-APR-2001 (first entry)
XX Human PRO270 protein.
XX
XX Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
XX antiparkinsonian; neurotropic; neuroprotective; vulnery; cardiant;
XX antiangiogenic; vasotrophic; antisthmatic; antirheumatic; cancer;
XX antiarthritic; antifertility; antidiabetic; antiviral; diabetes;
XX ophthalmologic; gene therapy; skin disease; gastrointestinal disorder;
XX ischaemia; inflammation.
XX
XX Homo sapiens.
```



```

XX PN W0200104311-AL.
XX PD 18-JAN-2001.
XX PF 22-FEB-2000; 2000WO-US04414.
XX PR 07-JUL-1999; 99US-0143048.
XX PR 26-JUL-1999; 99US-0145698.
XX PR 28-JUL-1999; 99US-0146222.
XX PR 08-SEP-1999; 99WO-US20594.
XX PR 13-SEP-1999; 99WO-US20944.
XX PR 15-SEP-1999; 99WO-US21090.
XX PR 15-SEP-1999; 99WO-US21547.
XX PR 05-OCT-1999; 99WO-US223089.
XX PR 29-NOV-1999; 99WO-US28214.
XX PR 30-NOV-1999; 99WO-US28313.
XX PR 16-DEC-1999; 99WO-US30095.
XX PR 20-DEC-1999; 99WO-US30911.
XX PR 20-DEC-1999; 99WO-US30999.
XX PR 05-JAN-2000; 99WO-US00219.
XX PA (GETH ) GENENTECH INC.
XX PI Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
XX PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
XX PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kljavin LJ;
XX PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
XX PI Williams PM, Wood WT;
XX WPI: 2001-081051/09.
XX DR N-PSDB; AAF72409.
XX PT Sixty one nucleic acids encoding PRO polypeptides which are useful in
XX PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
XX PT squamous cell carcinoma) and neurodegenerative diseases (e.g.
XX PT Alzheimer's disease) -
XX PS Claim 1; Fig 76; 393pp; English.
XX CC The present sequence is one of sixty one novel secreted and
XX CC transmembrane PRO polypeptides. The PRO polypeptides are
XX CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
XX CC squamous cell carcinoma), gastrointestinal disorders (e.g.
XX CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
XX CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.
XX CC endometrial bleeding angiogenesis, ischaemias such as coronary
XX CC ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,
XX CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
XX CC diabetes and retinal disorders such as retinitis pigmentosum.
XX CC The PRO nucleic acids have applications in molecular biology, including
XX CC use as hybridization probes, and in chromosome and gene mapping.
XX SQ Sequence 296 AA;

Query Match 97.6%; Score 1326; DB 22; Length 296;
Best Local Similarity 86.8%; Pred. No. 2.7e-136;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

QY 1 MAVLAPLIALVSVPRLSRWLAOPYYLLSALLSAFLLVRLKPLCHGLPTQREDGNPCD 60
DB 1 MAVLAPLIALVSVPRLSRWLAOPYYLLSALLSAFLLVRLKPLCHGLPTQREDGNPCD 60
QY 61 FQWREVEILMFLSAIYVMKNRRS-----83
DB 61 FQWREVEILMFLSAIYVMKNRRSITVEQHTGNFMFSKVANTILFFRLDMGLLYITLC 120
QY 84 -MFLMTCKPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWNCQSFAPYADL 142
DB 121 IVFLMTCKPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWNCQSFAPYADL 180
QY 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKVTSPITKQLPTLILFQGGKAMRRPOIDKKGR 202

```

```

DB 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKVTSPITKQLPTLILFQGGKAMRRPOIDKKGR 240
QY 203 AVSWTFSEENVIREFNLYORAKKLSKAGDNIPEEOPVASTPTTTSVDGENKKDK 258
DB 241 AVSWTFSEENVIREFNLYORAKKLSKAGDNIPEEOPVASTPTTTSVDGENKKDK 296

RESULT 6
ABBI2004
ID ABBI2004 standard; peptide; 329 AA.
XX AC
XX ABBI2004;
XX DT 11-JAN-2002 (first entry)
XX DE Human CG1-31 protein homologue, SEQ ID NO:2374.
XX KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antitumor.
XX OS Homo sapiens.
XX PN W0200157188-A2.
XX PD 09-AUG-2001.
XX PF 05-FEB-2001; 2001WO-US03800.
XX PR 03-FEB-2000; 2000US-0496914.
XX PR 27-APR-2000; 2000US-0560875.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX WPI: 2001-457740/49.
XX DR N-PSDB; ABA09248.
XX PT Human proteins and DNA encoding sequences useful for preventing,
XX PT treating or ameliorating a medical condition in a mammalian subject
XX PT e.g. arthritis and cancer -
XX PS Claim 20; Page 294-295; 1963pp; English.
XX CC Sequences ABBI0981-ABBI2330 represent 1350 novel human polypeptides, and
XX CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
XX CC invention also relates to vectors and recombinant host cells comprising a
XX CC nucleotide of the invention, methods of producing the novel polypeptides,
XX CC antibodies against the polypeptides, methods of detecting the nucleotides
XX CC or polypeptides in a sample, and methods of identifying compounds which
XX CC bind to polypeptides of the invention. Although novel, many of the
XX CC polypeptides of the invention have homology to known proteins, thereby
XX CC giving an insight into their probable biological activities, and hence
XX CC potential therapeutic applications. The polypeptides of the invention may
XX CC have various activities, including cytokine, cell proliferation or cell
XX CC differentiation activities; stem cell growth factor activity;
XX CC haematopoiesis regulatory activity; tissue growth activity;
XX CC immunomodulatory activity; activin- or inhibin-related activities;
XX CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
XX CC thrombolytic activities; receptor or ligand activities; or may be
XX CC involved in oncogenesis, cancer cell proliferation or metastasis.
XX CC depending on their biological activities, polypeptides and nucleotides of

```

CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.

XX Sequence 329 AA;

Query Match 97.6%; Score 1326; DB 22; Length 329;
 Best Local Similarity 86.8%; Pred. No. 3.2e-136;
 Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;
 QY 1 MAVLAPLIAVSVPRLSRWLAQPYLLSALLSAAFLVRLKPLPLCHGLPTQREDGNPCD 60
 DB 34 MAVLAPLIAVSVPRLSRWLAQPYLLSALLSAAFLVRLKPLPLCHGLPTQREDGNPCD 93
 QY 61 FDMREVEILMFLSAIVMKNRRS----- 83
 DB 94 FDMREVEILMFLSAIVMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 153
 QY 84 -MFLMTCRPLLYMGPEYIKYFNNDKTIDEELERDKRVTWVEFFANWSDCQSFAPYADL 142
 DB 154 IVFLMTCRPLLYMGPEYIKYFNNDKTIDEELERDKRVTWVEFFANWSDCQSFAPYADL 213
 QY 143 SLKYNCTGLNFGKVDVGRYTVSTRYKYSTPLTKOLPTLLIFQGGKEAMRRPQIDKGR 202
 DB 214 SLKYNCTGLNFGKVDVGRYTVSTRYKYSTPLTKOLPTLLIFQGGKEAMRRPQIDKGR 273
 QY 203 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
 DB 274 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 329

RESULT 7

ID AAW75204 standard; Protein; 296 AA.

XX AAW75204;

WT 29-JAN-1999 (first entry)

DE Human secreted protein encoded by gene 9 clone HJBCD89.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

XX Key Location/Qualifiers
 FT Misc-difference 281
 FT /label= unknown

PN WO9804083-A2.

XX

17-SEP-1998.

PD

12-MAR-1998; 98WO-US04858.

XX

19-DEC-1997; 97US-0068368.

PR

14-MAR-1997; 97US-0040710.

PR

30-MAY-1997; 97US-0048100.

PR

30-MAY-1997; 97US-0048189.

PR

30-MAY-1997; 97US-0048357.

PR

06-JUN-1997; 97US-0048970.

PR

03-SEP-1997; 97US-0057765.

XX

(HUMA-) HUMAN GENOME SCI INC.

PA

Ferrie AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;

PI

Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;

PI

Wei YF, Young PE, Zeng Z;

XX

WPI; 1998-520811/44.

DR

N-PSDB; AAV34294.

XX

Isolated human poly:nucleotide(s) encoding secretory peptide(s) -

PT

used to develop products for the diagnosis and treatment of e.g.

PT

inflammation, cancers, CNS disorders or immune system disorders

XX

Claim 1; Page 158; 201pp; English.

PS

XX This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AAV34277) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 28 novel genes and their fragments (nucleic
 CC acid sequences: AAV34286-V34325; amino acid sequences AAW75196-W75235)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 28
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAV34286 for described uses).

XX Sequence 296 AA;

Query Match 97.3%; Score 1322; DB 19; Length 296;
 Best Local Similarity 86.5%; Pred. No. 7.5e-136;

Matches 256; Conservative 1; Mismatches 1; Indels 38; Gaps 1;

QY 1 MAVLAPLIAVSVPRLSRWLAQPYLLSALLSAAFLVRLKPLPLCHGLPTQREDGNPCD 60

DB 1 MAVLAPLIAVSVPRLSRWLAQPYLLSALLSAAFLVRLKPLPLCHGLPTQREDGNPCD 60

QY 61 FDMREVEILMFLSAIVMKNRRS----- 83

DB 61 FDMREVEILMFLSAIVMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120

QY 84 -MFLMTCRPLLYMGPEYIKYFNNDKTIDEELERDKRVTWVEFFANWSDCQSFAPYADL 142

DB 121 IVFLMTCRPLLYMGPEYIKYFNNDKTIDEELERDKRVTWVEFFANWSDCQSFAPYADL 180

QY 143 SLKYNCTGLNFGKVDVGRYTVSTRYKYSTPLTKOLPTLLIFQGGKEAMRRPQIDKGR 202

DB 181 SLKYNCTGLNFGKVDVGRYTVSTRYKYSTPLTKOLPTLLIFQGGKEAMRRPQIDKGR 240

QY 203 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258

DB 241 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296

RESULT 8

AAW75226
ID AAW75226 standard; Protein; 293 AA.
XX
XX AC
XX AAW75226;
XX DT 29-JAN-1999 (first entry)
XX DE Human secreted protein encoded by gene 9 clone HJBCD89.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX Homo sapiens.
XX PN WO9840483-A2.
XX PD 17-SEP-1998.
XX PF 12-MAR-1998; 98WO-US04858.
XX PR 19-DEC-1997; 97US-0068368.
XX PR 14-MAR-1997; 97US-0040710.
XX PR 14-MAR-1997; 97US-0040762.
XX PR 30-MAY-1997; 97US-0048100.
XX PR 30-MAY-1997; 97US-0048189.
XX PR 30-MAY-1997; 97US-0048357.
XX PR 30-MAY-1997; 97US-0050934.
XX PR 06-JUN-1997; 97US-0048970.
XX PR 05-SEP-1997; 97US-0057765.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ferrie AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;
PI Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;
PI Wei YF, Young PE, Zeng Z;
XX WPI; 1998-520811/44.
XX DR N-PSDB; AAV34316.
XX
XX Isolated human poly:nucleotide(s) encoding secretory peptide(s) -
XX used to develop products for the diagnosis and treatment of e.g.
XX inflammation, cancers, CNS disorders or immune system disorders
XX
XX Claim 1: Page 173-174; 201pp; English.

XX This sequence represents a secreted human protein encoded by the gene
XX clone detailed in the descriptor line. The gene can be used to generate
XX fusion proteins by linking to the gene to a human immunoglobulin Fc
XX portion (e.g. AAV34277) for increasing the stability of the fused
XX protein as compared to the human protein only.
XX The invention relates to 28 novel genes and their fragments (nucleic
XX acid sequences: AAV34286-V34325; amino acid sequences AAW75196-W75235)
XX which are useful for preventing, treating or ameliorating medical
XX conditions e.g. by protein or gene therapy. Also, pathological
XX conditions can be diagnosed by determining the amount of the new
XX polypeptides in a sample or by determining the presence of mutations in
XX the new polynucleotides. Specific uses are described for each of the 28
XX polynucleotides, based on which tissues they are most highly expressed in
XX (see AAV34286 for described uses).

XX Sequence 293 AA;
XX Query Match 96.6%; Score 1313; DB 19; Length 293;
XX Best Local Similarity 86.7%; Pred. No. 7.1e-135;
XX Matches 254; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

QY 4 LAPLIALVSVPRLSRWLAOPYLLSALLSAAFLLVKRLPLCHGLPTQREDGNPCDFW 63

DB 1 LAPLIALVSVPRLSRWLAOPYLLSALLSAAFLLVKRLPLCHGLPTQREDGNPCDFW 60
QY 64 REVEILMFLSAIVMMKNRRS-----MF 85
DB 61 REVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVF 120
QY 86 LMTCKPPLYMGPEYIKYFNDKTIIDEELERDKRVTWIIEFFANWSNDCQSFAPYADLSLK 145
DB 121 LMTCKPPLYMGPEYIKYFNDKTIIDEELERDKRVTWIIEFFANWSNDCQSFAPYADLSLK 180
QY 146 YNCTGLNFGKVDVGRYTDVSTRYKVTSPKLTQPLTLILFOGGKEAMRRPOIDKKGRAVS 205
DB 181 YNCTGLNFGKVDVGRYTDVSTRYKVTSPKLTQPLTLILFOGGKEAMRRPOIDKKGRAVS 240
QY 206 WTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQPVASTPTTYSVDGENKKDK 258
DB 241 WTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQPVASTPTTYSVDGENKKDK 293

RESULT 9

ABB06300

ID ABB06300 standard; Protein; 296 AA.

XX AC ABB06300;

XX DT 01-JUN-2002 (first entry)

XX DE Prostaglandin E2 (PGE2) protein.

XX KW Prostaglandin E2; PGE2; stomach cancer.

XX OS Unidentified.

XX PN KR2001081233-A.

XX PD 29-AUG-2001.

XX PF 11-FEB-2000; 2000KR-0006416.

XX PR 11-FEB-2000; 2000KR-0006416.

XX PA (NAHE-) NAT INST HEALTH IN KOREA.

XX PI Kim GC, Nam MJ, Park MS;

XX DR WPI; 2002-136917/18.

XX DR N-PSDB; ABL49614.

XX Novel gene expressed by prostaglandin E2 (PGE2) in stomach cancer cell
XX line snu-1 and its amino acid sequence -
XX Disclosure; Page 12; 12pp; Korean.
XX The present sequence represents prostaglandin E2 (PGE2) which is
XX expressed in the stomach cancer cell line SNU-1. The PGE2 gene can be
XX used in the treatment of stomach cancer.

SQ Sequence 296 AA;

Query Match 96.2%; Score 1308; DB 23; Length 296;
Best Local Similarity 86.1%; Pred. No. 2.6e-134;
Matches 255; Conservative 1; Mismatches 2; Indels 38; Gaps 1;

QY 1 MAVLAPLIALVSVPRLSRWLAOPYLLSALLSAAFLLVKRLPLCHGLPTQREDGNPCD 60
DB 1 MAVLAPLIALVSVPRLSRWLAOPYLLSALLSAAFLLVKRLPLCHGLPTQREDGNPCD 60

QY 61 FDRWEVEILMFLSAIVMMKNRRS----- 83

DB 61 FDRWEVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120

QY 84 -MFLMTCPPPLYMGPEYIKYFNDKTIIDEELERDKRVTWIIEFFANWSNDCQSFAPYADL 142

Db 121 IIVFLMTCKPPLYMGPEYIKYFNDKTIIDEELERKRVTIWIEFFANWSDCQSFAPYADL 180
 QY 143 SLKYNTCTGLNFGKVDVGRYTDVSTRYKVTSPKLTQLPTLILFQGGKEAMRRPQIDKKGR 202
 Db 181 SLKYNTCTGLNFGKVDVGRYTDVSTRYKASTSPKLTQLPTLILFQGGKEAMRRPQIDKKGR 240
 QY 203 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
 Db 241 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296

RESULT 10

AAB88335
 ID AAB88335 standard; Protein; 372 AA.

XX AC AAB88335;

DT 23-MAY-2001 (first entry)

XX Human membrane or secretory protein clone PSEC0045.

DE Human; secretory protein; membrane protein; vaccine; gene therapy;
 vw rheumatoid arthritis; diabetes.

OS Homo sapiens.

XX EP1067182-A2.

PN 10-JAN-2001.

XX 07-JUL-2000; 2000EP-0114090.

XX 08-JUL-1999; 99JP-0194179.

PR 11-JAN-2000; 2000JP-0118775.

PR 02-MAY-2000; 2000JP-0183766.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;

XX WPI; 2001-093989/11.

XX N-PSDB; AAF93762.

XX Nucleic acids encoding secretory proteins/membrane proteins, useful in
 PT gene therapy or as candidate target molecules in drug development -

PS Claim 1; SEQ ID 38; 609pp + CD ROM; English.

XX This invention relates to nucleic acid sequences AAF93744 - AAF93916
 CC which encode human secretory or membrane proteins represented by
 CC AAB88317 - AAB88419. Included in the invention are primers
 CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
 CC cDNA sequences of the invention. The invention also includes methods for
 CC the production of antibodies directed against the proteins, and cDNA
 CC sequences, which can be used in vaccines. The polynucleotide sequences
 CC can be used in gene therapy. The polynucleotide sequences and the
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate secretory
 CC protein/membrane protein expression. The nucleic acids and complementary
 CC sequences may also be used as DNA probes in diagnostic assays
 CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples. They may also be
 CC used to study the expression and function of secretory proteins/membrane
 CC polypeptides and their role in metabolism. The polypeptides may be used
 CC as antigens in the production of antibodies against them and in assays to
 CC identify modulators (agonists and antagonists) of expression and
 CC activity. The antibodies and antagonists may also be used as therapeutic
 CC agents to down regulate expression and activity. The antibodies may also
 CC be used as diagnostic agents for detecting the presence of the
 CC polypeptides in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA). Examples of diseases which may be treated include rheumatoid
 CC arthritis and diabetes.

XX SQ Sequence 372 AA;
 Query Match 92.6%; Score 1259; DB 22; Length 372;
 Best Local Similarity 83.0%; Pred. No. 8.3e-129;
 Matches 244; Conservative 4; Mismatches 8; Indels 38; Gaps 1;
 QY 1 MAVLAPLIALIYVSPRLSRWLQAPYYLLSALLSAAFLLVKRLPLCHGLPTQREDGNPCD 60
 Db 1 MAVLAPLIALIYVSPRLSRWLQAPYYLLSALLSAAFLLVKRLPLCHGLPTQREDGNPCD 60
 QY 61 FDMREVEILMFLSAIVMKNNRS----- 83
 Db 61 FDMREVEILMFLSAIVMKNNRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
 QY 84 -MFLMTCKPPLYMGPEYIKYFNDKTIIDEELERKRVTIWIEFFANWSDCQSFAPYADL 142
 Db 121 IIVFLMTCKPPLYMGPEYIKYFNDKTIIDEELERKRVTIWIEFFANWSDCQSFAPYADL 180
 QY 143 SLKYNTCTGLNFGKVDVGRYTDVSTRYKVTSPKLTQLPTLILFQGGKEAMRRPQIDKKGR 202
 Db 181 SLKYNTCTGLNFGKVDVGRYTDVSTRYKVTSPKLTQLPTLILFQGGKEAMRRPQIDKKGR 240
 QY 203 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKK 256
 Db 241 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKK 294

RESULT 11

ABG18889
 ID ABG18889 standard; Protein; 344 AA.

XX AC ABG18889;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #18880.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0549167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS83076.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID No 49248; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving


```
XX Homo sapiens.
XX
XX WO98040483-A2.
XX
XX 17-SEP-1998.
XX
XX 12-MAR-1998; 98WO-US04858.
XX
XX 19-DEC-1997; 97US-0068368.
XX
XX 14-MAR-1997; 97US-0040710.
XX
XX 14-MAR-1997; 97US-0040762.
XX
XX 30-MAY-1997; 97US-0048100.
XX
XX 30-MAY-1997; 97US-0048189.
XX
XX 30-MAY-1997; 97US-0048357.
XX
XX 30-MAY-1997; 97US-0050934.
XX
XX 06-JUN-1997; 97US-0048970.
XX
XX 05-SEP-1997; 97US-0057765.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ferrie AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;
XX Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;
XX Wei YF, Young PE, Zeng Z;
XX
XX WPI: 1998-520811/44.
XX N-PSDB; AAV34294.
XX
XX Isolated human poly:nucleotide(s) encoding secretory peptide(s) -
XX used to develop products for the diagnosis and treatment of e.g.
XX inflammation, cancers, CNS disorders or immune system disorders
XX
XX Disclosure; Page 11; 201pp; English.
XX
XX This sequence represents a fragment of a secreted human protein encoded
XX by the nucleic acid molecule designated Gene 9 (AAV34294). The gene can
XX be used to generate fusion proteins by linking to the gene to a human
XX immunoglobulin Fc portion (e.g. AAV34277) for increasing the stability
XX of the fused protein as compared to the human protein only.
XX The invention relates to 28 novel genes and their fragments (nucleic
XX acid sequences: AAV34286-V34325; amino acid sequences AAV75196-W75235)
XX which are useful for preventing, treating or ameliorating medical
XX conditions e.g. by protein or gene therapy. Also, pathological
XX conditions can be diagnosed by determining the amount of the new
XX polypeptides in a sample or by determining the presence of mutations in
XX the new polynucleotides. Specific uses are described for each of the 28
XX polynucleotides, based on which tissues they are most highly expressed in
XX (see AAV34286 for described uses).
XX
XX Sequence 216 AA:
Query Match 67.4%; Score 916; DB 19; Length 216;
Best Local Similarity 81.9%; Pred. No. 1.2e-91;
Matches 177; Conservative 1; Mismatches 0; Indels 38; Gaps 1;
QY 55 DGNPCDFDREVEILMFLSAIVMMKNRRS----- 83
DB 1 DGNPCDFDREVEILMFLSAIVMMKNRRSITVEQHIGNIFMESKVANTILFFRLDIRMGL 60
QY 84 -----MFLMTCKPPLYMGPEYIKYFNDKTIDEELEDRKRVTWIVEFFANWSNDCQSFA 136
DB 61 LYITLCIVFLMTCRPPLYMGPEYIKYFNDKTIDEELEDRKRVTWIVEFFANWSNDCQSFA 120
QY 137 PIYADLSLKYNCTGLNFGKVDVGRYTDVSTRYKVSTSPITKOLPTLILFOGGEAMRRPQ 196
DB 121 PIYADLSLKYNCTGLNFGKVDVGRYTDVSTRYKVSTSPITKOLPTLILFOGGEAMRRPQ 180
QY 197 IDKKGRAVSWTFSEENVITREFNLNELYQRAKKLSKA 232
DB 181 IDKKGRAVSWTFSEENVITREFNLNELYQRAKKLSKA 216
XX
XX RESULT 14.
```

```
AA01002
ID AAG01002 standard; Protein; 111 AA.
XX
XX AAG01002;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein, SEQ ID NO: 5083.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI: 2000-500381/45.
XX N-PSDB; AAC01008.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 13; SEQ ID 5083; 71pp + CD-ROM; English.
XX
XX The present sequence is a polypeptide encoded by one of a large number
XX of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
XX were prepared from total human RNAs or polyA+ RNAs derived from 30
XX different tissues. EST sequences usually correspond mainly to the 3'
XX untranslated region (UTR) of the mRNA because they are often obtained
XX from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
XX isolating cDNA sequences derived from the 5' ends of mRNAs and even in
XX those cases where longer cDNA sequences have been obtained, the full 5'
XX UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
XX ends and can therefore be used to obtain full length cDNAs and genomic
XX DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
XX chromosome mapping procedures. They are used to obtain upstream
XX regulatory sequences and to design expression and secretion vectors.
XX
XX Sequence 111 AA;
Query Match 39.4%; Score 536; DB 21; Length 111;
Best Local Similarity 98.0%; Pred. No. 1.7e-50;
Matches 100; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 95 MGPEYIKYFNDKTIDEELEDRKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNCTGLNFG 154
DB 10 VGPEYIKYFNDKTIDEELEDRKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNCTGLNFG 69
QY 155 KYDVGRTDVSTRYKVSTSPITKOLPTLILFOGGEAMRRPQ 196
DB 70 KYDVGRTDVSTRYKVSTSPITKOLPTLILFOGGEAMRRPQ 111
XX
XX RESULT 15
XX AAY12372
ID AAY12372 standard; Protein; 144 AA.
XX
XX AAY12372;
XX
XX 17-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO:403.
```

```

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS Homo sapiens.
XX
PN WO9906548-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-TB01222.
XX
OI-AUG-1997; 97US-0905135.
XX
(AA ) GENSET.
XX
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
DR WPI: 1999-153778/13.
XX
DR N-PSDB; AAX41205.
XX
PT New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
PT kidney, lung, umbilical cord, placenta and colon tissue
XX
PS Claim 27; Page 725; 824pp; English.
XX
CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY12261 to
CC AAY12514, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/ chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX
SQ Sequence 144 AA;
Query Match 37.8%; Score 514; DB 20; Length 144;
Best Local Similarity 71.5%; Pred. No. 6.6e-48;
Matches 103; Conservative 1; Mismatches 2; Indels 38; Gaps 1;

QY 1 MAVLAPLIAIVSVPLSRWLAQPYLLSALLSAFLVLRKLPPLCHGLPTQREDGNPCD 60
Db 1 MAVLAPLIAIVSVPLSRWLAQPYLLSALLSAFLVLRKLPPLCHGLPTQREDGNXCD 60
QY 61 FDWREVEILMFLSAIVMMKNRRS----- 83
Db 61 FDWREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLXITLC 120
QY 84 -MFLMTCRPPLYMGPEYIKYFNDK 106
Db 121 IVFLMTCRPPLYMGPEYIIXYFNDK 144

```

Search completed: July 9, 2003, 11:55:30
Job time : 75 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1520	97.7	1527	21	AAAF15991		Human protein cion
2	1511	97.2	1598	21	AAAF16316		Human prostate can
3	1429	91.9	1730	22	ABA09248		Human GGI-31 prote
4	1381.6	88.8	1699	19	AAV34294		Human secreted pro
5	1370.2	88.1	1631	22	AAF93782		Human cDNA encodin
6	1370	88.1	1620	22	AAS45940		Human DNA encoding
7	1368	88.0	1618	20	AAW52251		Protein PRO270 cDN
8	1361.8	87.6	1652	19	AAV34316		Human secreted pro
9	1361.2	87.5	1659	24	ABL49614		Prostaglandin E2 (

PR 25-AUG-1998; 98JP-0238116.
PR 09-SEP-1998; 98JP-0254736.
PR 29-SEP-1998; 98JP-0275505.
XX
PA (SAGA) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX
PI Kato S, Kimura T;
PI
XX
DR WPI; 2000-182694/16.
DR P-PSDB; AAY94893.
XX
PT Novel human proteins having hydrophobic domains useful for treating
PT osteoporosis, Alzheimer's disease, Parkinson's disease, asthma,
PT multiple sclerosis, rheumatoid arthritis, cancer, anaemia, and stroke -
XX
PS Claim 4; Page 341-343; 351pp; English.
XX
CC This sequence encodes a human protein of the invention, which has
CC hydrophobic domains. The DNA sequences can be used as a probe or as a
CC genetic marker. The protein can also be used as a marker, and to identify
CC potential genetic disorders. The DNA and protein can also be used as
CC nutritional sources or supplements. The protein exhibits cytokine, cell
CC proliferation, cell differentiation activities and induces production of
CC other cytokines in certain cell populations. The protein also exhibits
CC immune stimulating or immune suppressing activity. It can be used in the
CC treatment of various immune deficiencies and disorders, and to treat
CC infectious diseases caused by viral, bacterial, fungal or other
CC infections. The protein is also used for treating autoimmune disorders
CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid
CC arthritis. It is also useful in the treatment of allergic reactions and
CC conditions such as asthma, and in immune suppression after organ
CC transplantation. The protein is useful in regulation of haematopoiesis
CC and consequently in the treatment of myeloid or lymphoid cell
CC deficiencies. It is also used in compositions for tissue growth or
CC regeneration. The protein is also used in the treatment of osteoporosis
CC or osteoarthritis and in the treatment of periodontal disease and other
CC tooth repair processes. The protein is used in the treatment of nervous
CC system disorders such as Alzheimer's disease, Parkinson's disease, and
CC Huntington's disease. They are useful for protection or regeneration and
CC treatment of lung or liver fibrosis, reperfusion injury in various
CC tissues, and conditions resulting from systemic cytokine damage. They are
CC also used for promoting or inhibiting tissue differentiation. They are
CC also used as contraceptives since they exhibit activin or inhibin related
CC activities and as a fertility inducing therapeutic. They are used for
CC treating various coagulation disorders and in treatment and prevention of
CC conditions resulting from coagulation activities e.g. myocardial
CC infarction or stroke. They also act as receptors, receptor ligands or
CC inhibitors or agonists of receptor/ligand interactions. They are used to
CC treat inflammatory conditions such as septic shock, sepsis, ischaemia
CC reperfusion injury, arthritis, and nephritis. They can be used to
CC prevent tumours.

Sequence 1527 BP; 382 A; 373 C; 361 G; 411 T; 0 other;
Query Match 97.7%; Score 1520; DB 21; Length 1527;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GCAGTGGCCGTTAGCGCGCGGAAAGATGGCGTCTTGGCACCTCTAAATGCTCTCGTGTAT 94
DB 1 GCAGTGGCCGTTAGCGCGCGGAAAGATGGCGTCTTGGCACCTCTAAATGCTCTCGTGTAT 60
QY 95 TCGGTGCGCGGATTTCCAGCATGGCTCGCCCAACTTACTACCTTCTCTGCGGCGCTGCTC 154
DB 61 TCGGTGCGCGGATTTCCAGCATGGCTCGCCCAACTTACTACCTTCTCTGCGGCGCTGCTC 120
QY 155 TCTGCTGCT 214
DB 121 TCTGCTGCT 180
QY 215 CGCGAAGACGGTACCCGCTGTGACTTTGACTGGAGAGAGTGGAGATCTCATGTTCTC 274

DB 181 CGCGAAGACGGTACCCGCTGTGACTTTGACTGGAGAGAAGTGGAGATCCTGATGTTTCTC 240
QY 275 AGTGCCATTGTGATGATGAAGAACCGAGATCCATGTTCTGTAGTACAGTGCAGAACCCCCC 334
DB 241 AGTGCCATTGTGATGATGAAGAACCGAGATCCATGTTCTGTAGTACAGTGCAGAACCCCCC 300
QY 335 CTATATATGGGCGCTGAGTATATCAAGTACTTCAATGATATAAACCATTGATGAGGAATTA 394
DB 301 CTATATATGGGCGCTGAGTATATCAAGTACTTCAATGATATAAACCATTGATGAGGAATTA 360
QY 395 GAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGGCAATTGGTCTTAATGACTGC 454
DB 361 GAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGGCAATTGGTCTTAATGACTGC 420
QY 455 CAATCATTTGGCCCTATCTATGCTGACCTCTCCCTTAAATACAACTGTACAGGGCTAAAT 514
DB 421 CAATCATTTGGCCCTATCTATGCTGACCTCTCCCTTAAATACAACTGTACAGGGCTAAAT 480
QY 515 TTTGGGAAGTGGATGTTGGACGCTATATCTGATGTTAGTACGCGTACAAAGTGAAGCACA 574
DB 481 TTTGGGAAGTGGATGTTGGACGCTATATCTGATGTTAGTACGCGTACAAAGTGAAGCACA 540
QY 575 TCACCCCTCACCAAGCAACTCCCTACCTGATCTCTTCCAAAGTGGCAAGGAGCAATG 634
DB 541 TCACCCCTCACCAAGCAACTCCCTACCTGATCTCTTCCAAAGTGGCAAGGAGCAATG 600
QY 635 CGGCGGCACAGATTGACAAGAGGCGGCTGTCTCATGAGACCTTCTCTGAGGAGAT 694
DB 601 CGGCGGCACAGATTGACAAGAGGCGGCTGTCTCATGAGACCTTCTCTGAGGAGAT 660
QY 695 GTGATCCGAGATTTAACTTAAATGAGCTATACAGCGGCGGCAAGAACTATCAAGGCT 754
DB 661 GTGATCCGAGATTTAACTTAAATGAGCTATACAGCGGCGGCAAGAACTATCAAGGCT 720
QY 755 GGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCGCCACACAGTGTGAGATGG 814
DB 721 GGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCGCCACACAGTGTGAGATGG 780
QY 815 GAAACACAGAGGATAAATAGATCCTCAGTGTGGCAGTGTCTCTCTCTCTCTCTCTCTCT 874
DB 781 GAAACACAGAGGATAAATAGATCCTCAGTGTGGCAGTGTCTCTCTCTCTCTCTCTCTCT 840
QY 875 AGGCTCTTTTCCATAACCAAGCCTGAGGCTGACGCTTTTATTTATTTATTTATTTATTT 934
DB 841 AGGCTCTTTTCCATAACCAAGCCTGAGGCTGACGCTTTTATTTATTTATTTATTTATTT 900
QY 935 CTGTGACTGGTGGGCGCAGCATGACGCTTCTGATTTTAAAGAGGAGCATCTAGGAAATGTC 994
DB 901 CTGTGACTGGTGGGCGCAGCATGACGCTTCTGATTTTAAAGAGGAGCATCTAGGAAATGTC 960
QY 995 AGGCACCTCAGAGAGGCGCTGCCATGCTGTGGCCCACTGTTTCACTGGAGCAAGAAAGA 1054
DB 961 AGGCACCTCAGAGAGGCGCTGCCATGCTGTGGCCCACTGTTTCACTGGAGCAAGAAAGA 1020
QY 1055 GATCTCATGAGCAGCGAGGGGAAATGTTTCCCTCCAAGCTTGGGTGAGTGTGTTAACTG 1114
DB 1021 GATCTCATGAGCAGCGAGGGGAAATGTTTCCCTCCAAGCTTGGGTGAGTGTGTTAACTG 1080
QY 1115 CTTATCAGCTATTTCAGACATCTCCATGTTTCTCCATGAAACTCTGTGGTTTCATCATTC 1174
DB 1081 CTTATCAGCTATTTCAGACATCTCCATGTTTCTCCATGAAACTCTGTGGTTTCATCATTC 1140
QY 1175 CTTCTTAGTGTACCTGCACAGCTTGGTGTAGACCTAGATTTAAACCTTAAGGTAAGTGTG 1234
DB 1141 CTTCTTAGTGTACCTGCACAGCTTGGTGTAGACCTAGATTTAAACCTTAAGGTAAGTGTG 1200
QY 1235 GGGTATAGAACGCTTAAGAAATTTTCCCCCAAGGACTCTTCTTAAAGCCCTTCTTGCT 1294
DB 1201 GGGTATAGAACGCTTAAGAAATTTTCCCCCAAGGACTCTTCTTAAAGCCCTTCTTGCT 1260
QY 1295 TCGTTTATGCTTCAATTTAAAGTATAGCCCTTAAGCTTTGTCGCTAGTCTTAAGGAGAAC 1354
DB 1261 TCGTTTATGCTTCAATTTAAAGTATAGCCCTTAAGCTTTGTCGCTAGTCTTAAGGAGAAC 1320

```
QY 1355 CTTTAAACCAAGTTTTTATCATTCAGACACATATTGACACCCCTATTTTGGGG 1414
DB 1321 CTTTAAACCAAGTTTTTATCATTCAGACACATATTGACACCCCTATTTTGGGG 1380
QY 1415 ATTGAGAAGGGTGAATAGAGGCTTGAGACTTTCCCTTTGTTGGTAGGACTTGGAGGAGA 1474
DB 1381 ATTGAGAAGGGTGAATAGAGGCTTGAGACTTTCCCTTTGTTGGTAGGACTTGGAGGAGA 1440
QY 1475 AATCCCTGGACTTTCACCTAACCTCTGACATACCTCCACACCCAGTTGATGGCTTTC 1534
DB 1441 AATCCCTGGACTTTCACCTAACCTCTGACATACCTCCACACCCAGTTGATGGCTTTC 1500
QY 1535 GTAATAAAAGATTGGGATT 1554
DB 1501 GTAATAAAAGATTGGGATT 1520

--RESULT 2
AAFI6316
ID AAFI6316 standard; cDNA; 1598 BP.
XX
AC
XX
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:751.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX
OS Homo sapiens.
XX
XX WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05988.
XX
PR 12-MAR-1999; 99US-0124270.
XX
..Y (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
-A Rosen CA, Ruben SW;
PI
XX WPI; 2000-587513/55.
-DR P-PSDB; AAB57113.
XX
XX Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
XX Claim 1; Page 1172; 2338pp; English.
XX
XX AAFI5566 to AAFI6505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
```


Qy	224	GGTAAACCCGCTGACATTTTGACTCGAGAGAAGTGGAGATCTCTGATGTTTCTCAGTGCACATTT	283
Db	181	GGTAAACCCGCTGACATTTTGACTCGAGAGAAGTGGAGATCTCTGATGTTTCTCAGTGCACATTT	240
Qy	284	GTGATGATGAAGAACCGCAGATCCA-----	308
Db	241	GTGATGATGAAGAACCGCAGATCCATCACTGTGGAGCAACATATAGSCAACATTTTTCATG	300
Qy	309	-----	308
Db	301	TTTAGTAAAGTGCCCAACACAATCTTTTCTTCGCTTGGATATTCGATGGGCTACTTT	360
Qy	309	-----TGTTCCCTGATCAGCTGCAACCCGCCCTTATATATGSGCCCT	349
Db	361	TACATCACACTCTGCGATATGTGTTTCCCTGATGACGTGCAACCCGCCCTTATATATGSGCCCT	420
Qy	350	GAGTATATCAAGTACTTCAATATGAATAAACCATTTGATGAGGAACATAGAACGGGACAAGAGG	409
Db	421	GAGTATATCAAGTACTTCAATATGAATAAACCATTTGATGAGGAACATAGAACGGGACAAGAGG	480
Qy	410	GTCACTTGGATTGTGGAGTTCTTTTGCCAATTGCTTAATGACTTGCCCAATCATTTTGCCCT	469
Db	481	GTCACTTGGATTGTGGAGTTCTTTTGCCAATTGCTTAATGACTTGCCCAATCATTTTGCCCT	540
Qy	470	ATCTATGCTGACCTCTCCCTTAAATACAACTGTACAGGGCTAAATTTTGGGAAGGTGGAT	529
Db	541	ATCTATGCTGACCTCTCCCTTAAATACAACTGTACAGGGCTAAATTTTGGGAAGGTGGAT	600
Qy	530	GTGTGGAGCGCTACTGATGTTAGTACGCGGTACAAAGTGAGCACATCACCCCTCACCAAG	589
Db	601	GTGTGGAGCGCTACTGATGTTAGTACGCGGTACAAAGTGAGCACATCACCCCTCACCAAG	660
Qy	590	CAACTCCCTACCCTGATCCTGTTTCCAAGGTGGCAAGGAGCAATGCGCGCGGCCACAGATT	649
Db	661	CAACTCCCTACCCTGATCCTGTTTCCAAGGTGGCAAGGAGCAATGCGCGCGGCCACAGATT	720
Qy	650	GACAAGAAAGGACGGGCTGTCTCATGACCTTCTCTGAGGAGATGTGATCCGAGAAATTT	709
Db	721	GACAAGAAAGGACGGGCTGTCTCATGACCTTCTCTGAGGAGATGTGATCCGAGAAATTT	780
Qy	710	AACTTAAATGAGCTATACCAGCGGGCCAAAGAACTATCAAAGGCTGGAGACAATATCCCT	769
Db	781	AACTTAAATGAGCTATACCAGCGGGCCAAAGAACTATCAAAGGCTGGAGACAATATCCCT	840
Qy	770	GAGGAGCAGCCTGTGGCTTCAACCCCCACACAGTGTGAGATGGGAAAAACAAGAAGGAT	829
Db	841	GAGGAGCAGCCTGTGGC - TCAACCCCCACACAGTGTGAGATGGGAAAAACAAGAAGGAT	899
Qy	830	AAATAAGATCCTCACTTTGGGAGTGCTTCTCTCCTGTCAATTCAGGCTCTTTTCATATA	889
Db	900	AAATAAGATCCTCACTTTGGGAGTGCTTCTCTCCTGTCAATTCAGGCTCTTTTCATATA	959
Qy	890	CCACAGCCTGAGGCTGACGCTTTATTTATGTTTTCCCTTTTCGCTGTGACTGGGTGGG	949
Db	960	CCACAGCCTGAGGCTGACGCTTTATTTATGTTTTCCCTTTTGCTGTGACTGGGTGGG	1019
Qy	950	GCAGCTGACGCTTCTGATTTTAAAGAGGCACTTAGGGAATGTGACGACCCCTACAGGA	1009
Db	1020	GCAGCTGACGCTTCTGATTTTAAAGAGGCACTTAGGGAATGTGACGACCCCTACAGGA	1079
Qy	1010	AGGCCTGCCATGTGTGGCAACTGTTTCACTGGAGCAAGAAAGAGATCTCATAGGAGG	1069
Db	1080	AGGCCTGCCATGTGTGGCAACTGTTTCACTGGAGCAAGAAAGAGATCTCATAGGAGG	1139
Qy	1070	AGGGGGAATGGTTCCCTCCAAGCTTGGGTCACTGTGTTAACTGCTATACAGCTATTCA	1129
Db	1140	AGGGGGAATGGTTTCCCTCCAAGCTTGGGTCACTGTGTTAACTGCTATACAGCTATTCA	1199
Qy	1130	GACATCCCATGGTTTCTCCATGAACCTCTGTGGTTTCATCATTCCTCTTAGTTGACCT	1189
Db	1200	GACATCCCATGGTTTCTCCATGAACCTCTGTGGTTTCATCATTCCTCTTAGTTGACCT	1259
Qy	1190	GCACAGCTTGTTTAGACTAGATTTAAACCTTAAGGTAAAGTGTGGGTATAGAACGCTA	1249

Db	1260	GCACGCTTTGGTTAGACCTAGATTTAACCTTAAGGTAAAGTCTGGGGTATAGAACCCTA	1319
Qy	1250	AGAAATTTTCCCCCAAGGACTCTTGCTTCCCTTAAGCCCTTCTGGCTTCGTTATATGGTCTTC	1309
Db	1320	AGAATTTTCCCCCAAGGACTCTTGCTTCCCTTAAGCCCTTCTGGCTTCGTTATATGGTCTTC	1379
Qy	1310	ATTTAAAGTATAAGCCTTAACCTTGTGCTCCTAGTCTTAAGGAGAAACCTTTTAACCAAAAGT	1369
Db	1380	ATTTAAAGTATAAGCCTTAACCTTGTGCTCCTAGTCTTAAGGAGAAACCTTTTAACCAAAAGT	1439
Qy	1370	TTTTATCATTTGAAGACAATATTGAACAACCCCTATTTTGTGGGGATTGAGAAGGGGTGA	1429
Db	1440	TTTTATCATTTGAAGACAATATTGAACAACCCCTATTTTGTGGGGATTGAGAAGGGGTGA	1499
Qy	1430	ATAGAGCCTTGAGACTTTCTCTTGTGTGGTGTAGGACTTTGAGAGGAGAAATCCCTTGGACTTT	1489
Db	1500	ATAGAGCCTTGAGACTTTCTCTTGTGTGGTGTAGGACTTTGAGAGGAGAAATCCCTTGGACTTT	1559
Qy	1490	CACCTAACCCCTCTGACATACCTCCGCACACCAGTTGATGGCTTTCCGCTTAATAAAAAAGATTG	1549
Db	1560	CACCTAACTCTGNCATACCTCCGCACACCAGTTGATGGCTTTCCGCTTAATAAAAAAGATTG	1619
	1550	GGATT	1554
Db	1620	GGATT	1624

RESULT 6	
AAS45940	
ID	AAS45940 standard; cDNA; 1620 BP.
XX	
XX	
AC	AAS45940;
XX	
XX	18-DEC-2001 (first entry)
DT	
XX	
XX	Human DNA encoding PRO polypeptide sequence #16.
DE	
XX	
KW	PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss;
KW	dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW	blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW	adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
XX	PCR primer.
XX	
OS	Homo sapiens.
XX	
XX	WO200168848-A2.
PN	
XX	
XX	
PD	20-SEP-2001.
XX	
XX	28-FEB-2001; 2001WO-US06520.
PF	

01-MAR-2000;	2000WO-US05601.
02-MAR-2000;	2000WO-US05841.
03-MAR-2000;	2000US-187202P.
06-MAR-2000;	2000US-186968P.
14-MAR-2000;	2000US-189320P.
14-MAR-2000;	2000US-189328P.
15-MAR-2000;	2000WO-US06884.
PR 21-MAR-2000;	2000US-190828P.
PR 21-MAR-2000;	2000US-191007P.
PR 21-MAR-2000;	2000US-191048P.
PR 28-MAR-2000;	2000US-191314P.
PR 29-MAR-2000;	2000US-192655P.
PR 29-MAR-2000;	2000US-193032P.
PR 29-MAR-2000;	2000US-193053P.
PR 30-MAR-2000;	2000WO-US08439.
PR 04-APR-2000;	2000US-194449P.
PR 04-APR-2000;	2000US-194647P.
PR 11-APR-2000;	2000US-195975P.
PR 11-APR-2000;	2000US-196000P.
PR 11-APR-2000;	2000US-196187P.
PR 11-APR-2000;	2000US-196690P.

PR 11-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX
XX
(GETH) GENENTECH INC.
PA
PA
XX
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
XX
WPI; 2001-602746/68.
P-PSDB: AAU29039.
DR
DR

Novel nucleic acids encoding PRO polypeptides, used to diagnose the presence of tumours, such as prostate and breast tumours, in mammals and to screen for modulators of the compounds -

Claim 2; Fig 31; 774pp; English.

Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR primers for PRO polypeptides of the invention. The sequences of the invention can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders.

SQ Sequence 1620 BP; 407 A; 394 C; 373 G; 442 T; 4 other;
 Query Match 88.1%; Score 1370; DB 22; Length 1620;
 Best Local Similarity 92.7%; Pred. No. 0;
 Matches 1494; Conservative 0; Mismatches 4; Indels 114; Gaps 1;

57	AGATGGCGGTCTTTGGCACCTCTAATTTGCTCTCGTGATTCGGTGCCGCGACTTTCACGAT	116
1	AGATGGCGGTCTTTGGCACCTCTAATTTGCTCTCGTGATTCGGTGCCGCGACTTTCACGAT	60
117	GGTCGCCCAACCTTACTACCTCTCTCGGCCCTGCTCTCTGCTGCTTCTCCTACTCGTGA	176
61	GGCTCGCCCAACCTTACTACCTCTCTCGGCCCTGCTCTCTGCTGCTTCTCCTACTCGTGA	120
177	GGAAATGCGCGCGCTCTGCCACGGTCTGCCACCCCAACGCGGAAGACGGTAAACCCGTGTG	236
121	GGAAATGCGCGCGCTCTGCCACGGTCTGCCACCCCAACGCGGAAGACGGTAAACCCGTGTG	180
237	ACTTTGACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCTATGTGATGATGAAGA	296
181	ACTTTGACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCTATGTGATGATGAAGA	240
297	ACCGCAGATCCA-----	308
OY		

|||||
Db 241 ACCGAGATCCATCACTGTGGAGCAACATATAGGCARCAATTTTCATGTTTAGTAAAGTGG 300
Qy 309 ----- 308
Db 301 CCAACACAATTTCTTCTTCGGCTTGGATATTGCGCATGGGCTACTTTACATCACACTCT 360
Qy 309 -----TGTTCCTGATGACGTGCAACCCCTTATATATGGCCCTCAGTATATCAAGT 362
Db 361 GCATAGTGTTCCTGATGACGTGCAACCCCTTATATATGGCCCTCAGTATATCAAGT 420
Qy 363 ACTTCATGATAAACCAATTCATGAGGAACATAGAACGGGCAAGAGGCTCACTTGGATTG 422
Db 421 ACTTCATGATAAACCAATTCATGAGGAACATAGAACGGGCAAGAGGCTCACTTGGATTG 480
Qy 423 TGGAGTTCTTTGGCAATTGGTCTTAATGACTGCCAATCATTTTGGCCCTATCTATGCTGACC 482
Db 481 TGGAGTTCTTTGGCAATTGGTCTTAATGACTGCCAATCATTTTGGCCCTATCTATGCTGACC 540
Qy 483 TCTCCCTTAATACAACCTGTACAGGGCTAAATTTTGGGAAGGTGATGTTGGACGCTATA 542
Db 541 TCTCCCTTAATACAACCTGTACAGGGCTAAATTTTGGGAAGGTGATGTTGGACGCTATA 600
Qy 543 CTGATGTTAGTACGGGTACAAAGTGTAGCATACACCCCTCACCAAGCAACTTCCCTACCC 602
Db 601 CTGATGTTAGTACGGGTACAAAGTGTAGCATACACCCCTCACCAAGCAACTTCCCTACCC 660
Qy 603 TGATCCTGTTCCAAAGGTGGCAAGGCAATGCGCGGCCACAGATTGACAAAGAAAGGAC 662
Db 661 TGATCCTGTTCCAAAGGTGGCAAGGCAATGCGCGGCCACAGATTGACAAAGAAAGGAC 720
Qy 663 GGGCTGCTCATGACACCTTCTCTGAGGAGATGTGATCGGAGATTTTAACTTAATGAGC 722
Db 721 GGGCTGCTCATGACACCTTCTCTGAGGAGATGTGATCGGAGATTTTAACTTAATGAGC 780
Qy 723 TATACCAGCGGCAAGAACTATCAAGGCTGGAGACAATATCCTCTGAGGAGCAGCCTG 782
Db 781 TATACCAGCGGCAAGAACTATCAAGGCTGGAGACAATATCCTCTGAGGAGCAGCCTG 840
Qy 783 TGGCTTCAACCCCAACCCACAGTGTGATGGGAAACAAAGAGGATAAATAAGATCCTC 842
Db 841 TGGCTTCAACCCCAACCCACAGTGTGATGGGAAACAAAGAGGATAAATAAGATCCTC 900
Qy 843 ACTTTGGCAGTGTCTCTCTCTCTGTCATTCAGGCTCTTTCCATACCAACCAAGCCTGAG 902
Db 901 ACTTTGGCAGTGTCTCTCTCTCTGTCATTCAGGCTCTTTCCATACCAACCAAGCCTGAG 960
Qy 903 GCTCAGCCTTTTATTTATGTTTCCCTTTGGCTGTGACTGGTGGGCGCAGCATCAGCT 962
Db 961 GCTCAGCCTTTTATTTATGTTTCCCTTTGGCTGTGACTGGTGGGCGCAGCATCAGCT 1020
Qy 963 TCTGATTTTAAAGAGGCATCTAGGGAATGTGAGGACCCCTACAGGAAGGCTGCGCATGC 1022
Db 1021 TCTGATTTTAAAGAGGCATCTAGGGAATGTGAGGACCCCTACAGGAAGGCTGCGCATGC 1080
Qy 1023 TGTGGCCAACTGTTTCACTGGAGCAAGAAAGATCTCATAGGAGGAGGGGGAATGTT 1082
Db 1081 TGTGGCCAACTGTTTCACTGGAGCAAGAAAGATCTCATAGGAGGAGGGGGAATGTT 1140
Qy 1083 TTCCCTCCAAGCTGGGTCAGTGTCTTAACCTTATCAGCTTATCAGCATCTCCATGG 1142
Db 1141 TTCCCTCCAAGCTGGGTCAGTGTCTTAACCTTATCAGCTTATCAGCATCTCCATGG 1200
Qy 1143 TTTCTCCATGAACCTCTGTGTTTTCATCTCTTCTTGTAGTTGACCTGCACAGCTTGGTT 1202
Db 1201 TTTCTCCATGAACCTCTGTGTTTTCATCTCTTCTTGTAGTTGACCTGCACAGCTTGGTT 1260
Qy 1203 AGACCTAGATTAAACCCCTAAGTAAGATGCTGGGGTATAGAACGCTAAGAAATTTTCCCCC 1262
Db 1261 AGACCTAGATTAAACCCCTAAGTAAGATGCTGGGGTATAGAACGCTAAGAAATTTTCCCCC 1320
Qy 1263 AAGGACTCTGCTTCAAGCCCTTCTGGCTTATGCTTATGCTCTTCAATAAAGTATAA 1322
|||||

1321 AAGGACTCTGCTTCTCTTAAGCCCTTCGGCTTCCTTTATGCTCTTCATTAAAAAGTATAA 1380
Qy 1323 GCCTAACCTTTGTCGTAGTCTTAAGGAGAAACCTTTAAACCACAAAGTTTTTATCATTTGAA 1382
Db 1381 GCCTAACCTTTGTCGTAGTCTTAAGGAGAAACCTTTAAACCACAAAGTTTTTATCATTTGAA 1440
Qy 1383 GACAAATATTGAACAAACCCCTTATTTTGTGGGATTGAGAAAGGGTGAATAGAGGCTTGAG 1442
Db 1441 GACAAATATTGAACAAACCCCTTATTTTGTGGGATTGAGAAAGGGTGAATAGAGGCTTGAG 1500
Qy 1443 ACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCTGGACITTTCACTAACCCCTCG 1502
Db 1501 ACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCTGGACITTTCACTAACCCCTCG 1560
Qy 1503 ACATACTCCCAACCCAGTTGATGCTTTCGCTAATAAAAAAGATTGGGATT 1554
Db 1561 ACATACTCCCAACCCAGTTGATGCTTTCGCTAATAAAAAAGATTGGGATT 1612

RESULT 7
AXS52251
ID AAX52251 standard; DNA; 1618 BP.
XX
AC AAX52251;
XX
XX 25-JUN-1999 (first entry)
XX
DE Protein PR0270 cDNA clone DNA39510-1181.
KW Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; AUS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophila areata;
KW anti-thrombotic; wound healing; tissue repair; ss.
XX Homo sapiens.
XX WO9914328-A2.
XX PD 25-MAR-1999.
XX
PF 16-SEP-1998; 98WO-US19330.
XX
PR 25-NOV-1997; 97US-0066840.
PR 17-SEP-1997; 97US-0059113.
PR 17-SEP-1997; 97US-0059115.
PR 17-SEP-1997; 97US-0059117.
PR 17-SEP-1997; 97US-0059119.
PR 17-SEP-1997; 97US-0059121.
PR 17-SEP-1997; 97US-0059122.
PR 17-SEP-1997; 97US-0059184.
PR 18-SEP-1997; 97US-0059263.
PR 18-SEP-1997; 97US-0059266.
PR 15-OCT-1997; 97US-0062125.
PR 17-OCT-1997; 97US-0062285.
PR 21-OCT-1997; 97US-0063486.
PR 24-OCT-1997; 97US-0062814.
PR 24-OCT-1997; 97US-0062816.
PR 24-OCT-1997; 97US-0063045.
PR 24-OCT-1997; 97US-0063120.
PR 24-OCT-1997; 97US-0063121.
PR 24-OCT-1997; 97US-0063127.
PR 24-OCT-1997; 97US-0063128.
PR 27-OCT-1997; 97US-0063329.
PR 27-OCT-1997; 97US-0063327.
PR 28-OCT-1997; 97US-0063541.
PR 28-OCT-1997; 97US-0063542.
PR 28-OCT-1997; 97US-0063544.
PR 28-OCT-1997; 97US-0063549.
PR 28-OCT-1997; 97US-0063550.

PR 28-OCT-1997; 97US-0063564.
PR 29-OCT-1997; 97US-0063435.
PR 29-OCT-1997; 97US-0063704.
PR 29-OCT-1997; 97US-0063732.
PR 29-OCT-1997; 97US-0063738.
PR 29-OCT-1997; 97US-0063734.
PR 29-OCT-1997; 97US-0064215.
PR 29-OCT-1997; 97US-0063735.
PR 31-OCT-1997; 97US-0063870.
PR 31-OCT-1997; 97US-0064103.
PR 03-NOV-1997; 97US-0064248.
PR 07-NOV-1997; 97US-0064809.
PR 12-NOV-1997; 97US-0065186.
PR 17-NOV-1997; 97US-0065846.
PR 18-NOV-1997; 97US-0065693.
PR 21-NOV-1997; 97US-0066120.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066772.
PR 24-NOV-1997; 97US-0066466.
PR 24-NOV-1997; 97US-0066770.
PR 24-NOV-1997; 97US-0066511.
PR 24-NOV-1997; 97US-0066453.
vX (GETH) GENENTECH INC.

PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
XX WPI: 1999-229533/19.
DR P-PSDB: AAY13380.
XX
PT New isolated human genes and polypeptides used in, e.g. treatment of
PT gastrointestinal ulceration
XX
PS Claim 2: Fig 75; 320pp; English.
XX
CC AAX5213-74 encode secreted and transmembrane human proteins, and are
CC obtained from cDNA libraries, prepared from fetal lung, fetal kidney,
CC fetal brain, fetal liver and fetal retina. The encoded polypeptides
CC have specific uses based on their homology to known polypeptides,
CC e.g. PRO211 and PRO217 can be used for disorders associated with the
CC preservation and maintenance of gastrointestinal mucosa and the repair
CC of acute and chronic mucosal lesions (e.g. enterocolitis,
CC Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital
CC microvillus atrophy), skin diseases associated with abnormal
CC keratinocyte differentiation (e.g. psoriasis, epithelial cancers such as
CC lung squamous cell carcinoma of the vulva and gliomas), potent effects on
CC cell growth and development, diseases related to growth or survival of
CC nerve cells including Parkinson's disease, Alzheimer's disease, ALS,
CC neuropathies or cancer. PRO265 can be used as for fibromodulin, e.g. for
CC reducing dermal scarring. PRO264 can be used as a target for anti-tumor
CC drugs. PRO533 may be used in the treatment of Usher Syndrome or Atrophila
CC areata; PRO269 can be used as an anti-thrombotic agent; PRO287
CC polypeptides and portions may have therapeutic applications in wound
CC healing and tissue repair; PRO317 can be used for treating problems of
CC the kidney, uterus, endometrium, blood vessels, or related tissue, e.g.
CC in the heart of genital tract.
XX
SQ Sequence 1618 BP; 406 A; 394 C; 372 G; 442 T; 4 other:

Query Match 88.0%; Score 1368; DB 20; Length 1618;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1492; Conservative 0; Mismatches 4; Indels 114; Gaps 1;

Qy 59 ATGGCGGTCTTGGACCTCTAATTCCTCTCGTATTCGGTCCCGGACTTTCAGATGG 118
Dy 1 ATGGCGGTCTTGGACCTCTAATTCCTCTCGTATTCGGTCCCGGACTTTCAGATGG 60

Qy 119 CTCGCCCAACCTTACTACTCTGTGCGCCCTGCTCTCTGCTGCTTCTCTACTCTGAGG 178
Dy 61 CTCGCCCAACCTTACTACTCTGTGCGCCCTGCTCTCTGCTGCTTCTCTACTCTGAGG 120

Qy 179 AACTGCGCGGCTTCTGCCAGGCTCTGCCACCCGAGAGCGGTAAACCCGTGAC 238
Dy 179 AACTGCGCGGCTTCTGCCAGGCTCTGCCACCCGAGAGCGGTAAACCCGTGAC 238

506 TTGTGGAGTCTTTGCCAATGGTCTAATGACTGCCAATCAATTTGCCCCATCTATCTGCTG 565
480 ACCTCTCCCTTAAATACAACTGTACAGGCTAAATTTGGGAAGTGGATGTTGGACGCT 539
566 ACCTCTCCCTTAAATACAACTGTACAGGCTAAATTTGGGAAGTGGATGTTGGACGCT 625
540 ATACTGATGTTAGTACGGGTACAAGTACAGACATCAACCCCTCACCAAGCAACTCCCTA 599
626 ATACTGATGTTAGTACGGGTACAAGTACAGACATCAACCCCTCACCAAGCAACTCCCTA 685
600 CCCTGATCTGTTCCTCAAGTGGCAAGGAGGCAATCGCGGGCCACACAGATTGCAAGAAAG 659
686 CCCTGATCTGTTCCTCAAGTGGCAAGGAGGCAATCGCGGGCCACACAGATTGCAAGAAAG 745
660 GACGGGCTGTCATGAGGCTTCTCTGAGGAGAAATGTATCGAGAAATTAACCTTAATG 719
746 GACGGGCTGTCATGAGGCTTCTCTGAGGAGAAATGTATCGAGAAATTAACCTTAATG 805
720 AGCTATACCGGGGCGCAAGAACTATCAAGGCTGGAGACATATCCCTGAGGAGCAGC 779
806 AGCTATACCGGGGCGCAAGAACTATCAAGGCTGGAGACATATCCCTGAGGAGCAGC 865
780 CTGTGGCTTCAACCCCCACACAGTGTACAGTGGGGAACCAAGAGGATTAATAAGATC 839
866 CTGTGGCTTCAACCCCCACACAGTGTACAGTGGGGAACCAAGAGGATTAATAAGATC 925
840 CTACTTTGGGAGTCTTCTCTCTCTGTCATTAATTCAGGCTCTTTCATACCAACCAAGCCT 899
926 CTACTTTGGGAGTCTTCTCTCTCTGTCATTAATTCAGGCTCTTTCATACCAACCAAGCCT 985
900 GAGGCTGAGCCCTTTATTTATGTTT - TCCCTTTGGCTGTGACTGGGTGGGGCAGCATGC 958
986 GAGGCTGAGCCCTTTATTTATGTTTCTCCCTTTGGCTGTGACTGGGTGGGGCAGCATGC 1045
959 AGCTTCTGATTTTAAAGAGGATCTAGGAAATGTTCAGGACCCCTACAGAGGCTGCTGCC 1018
1046 AGCTTCTGATTTTAAAGAGGATCTAGGAAATGTTCAGGACCCCTACAGAGGCTGCTGCC 1105
1019 ATGCTGTGCGCCAACTGTTTCACTGGAGCAAGAAAGATCTCATAGGAGGAGGGGAAA 1078
1106 ATGCTGTGCGCCAACTGTTTCACTGGAGCAAGAAAGATCTCATAGGAGGAGGGGAAA 1165
1079 TGGTTTCCCTTCAAGCTTGGGTCAGTGTGTTTAACTGCTTATCAGCTATTACAGATCTCC 1138
1166 TGGTTTCCCTTCAAGCTTGGGTCAGTGTGTTTAACTGCTTATCAGCTATTACAGATCTCC 1225
1139 ATGTTTCTCCATGAACCTCTGTGTTTTCATCATCTCTTCTAGTGAACCTGCACAGCTT 1198
1226 ATGTTTCTCCATGAACCTCTGTGTTTTCATCATCTCTTCTAGTGAACCTGCACAGCTT 1285
1199 GGTATAGACCTAGATTTAAACCTTAAGTAAAGATGCTGGGGTATAGAACGCTTAAGATTTTC 1258
1286 GGTATAGACCTAGATTTAAACCTTAAGTAAAGATGCTGGGGTATAGAACGCTTAAGATTTTC 1345
1259 CCCCAAGGACTCTGCTTCCCTTAAAGCCCTTGGCTTCTGTTATGCTCTTCAATAAAGT 1318
1346 CCCCAAGGACTCTGCTTCCCTTAAAGCCCTTGGCTTCTGTTATGCTCTTCAATAAAGT 1405
1319 ATAAGCCTAACTTTGTGCTAGTCTTAAAGGAGAAACCTTTAAACCAAAAGTTTATCAT 1378
1406 ATAAGCCTAACTTTGTGCTAGTCTTAAAGGAGAAACCTTTAAACCAAAAGTTTATCAT 1465
1379 TGAAGACATATTAAGAACACCCCTTATTTTGGGGATTGAGAGGGGTTGAATAGAGGCT 1438
1466 TGAAGACATATTAAGAACACCCCTTATTTTGGGGATTGAGAGGGGTTGAATAGAGGCT 1525
1439 TGAGACTTTCTTTGTGTTGTTAGGACTTGGAGGAGAAATCCCTTGGACTTTCTACTAACCC 1498
1526 TGAGACTTTCTTTGTGTTGTTAGGACTTGGAGGAGAAATCCCTTGGACTTTCTACTAACCC 1585
1499 TCTGACATACTCCCAACCCAGTTCATGCTTCCGTTAATAAAGATTTGGATT 1554
1586 TCTGACATACTCCCAACCCAGTTCATGCTTCCGTTAATAAAGATTTGGATT 1641

RESULT 10

AAF72409

ID : AAF72409 standard; cDNA; 1621 BP.

XX

AC AAF72409;

XX

DT 24-APR-2001 (first entry)

XX

DE Human PRO270 cDNA.

XX

Human: PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
antiParkinsonian nootropic; neuroprotective; vulnery; cardiant;
KW antiangiogenic; vasotropic; antiasthmatic; antineumatic; cancer;
KW antithratic; antinfertility; antidiabetic; antiviral; diabetes;
KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
ischaemia; inflammation; ss.

XX

OS Homo sapiens.

XX

PN WO200104311-A1.

XX

XX 18-JAN-2001.

XX

PF 22-FEB-2000; 2000WO-US04414.

XX

XX 07-JUL-1999; 99US-0143048.

XX

XX 26-JUL-1999; 99US-0145698.

XX

XX 28-JUL-1999; 99US-0146222.

XX

XX 08-SEP-1999; 99WO-US20594.

XX

XX 13-SEP-1999; 99WO-US20944.

XX

XX 15-SEP-1999; 99WO-US21090.

XX

XX 15-SEP-1999; 99WO-US21547.

XX

XX 05-OCT-1999; 99WO-US23089.

XX

XX 29-NOV-1999; 99WO-US28214.

XX

XX 30-NOV-1999; 99WO-US28313.

XX

XX 16-DEC-1999; 99WO-US30095.

XX

XX 20-DEC-1999; 99WO-US30911.

XX

XX 20-DEC-1999; 99WO-US30999.

XX

XX 05-JAN-2000; 99WO-US00219.

XX

(GETH) GENENTECH INC.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kljavin IJ;
Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
Williams PM, Wood WI;

WPI; 2001-081051/09.

P-PSDB; AAB80248.

Sixty one nucleic acids encoding PRO polypeptides which are useful in
the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
squamous cell carcinoma) and neurodegenerative diseases (e.g.
Alzheimer's disease)

Claim 2; Fig 75; 393pp; English.

The present sequence is one of sixty one nucleic acids encoding novel
secreted and transmembrane PRO polypeptides. The PRO polypeptides are
useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
squamous cell carcinoma), gastrointestinal disorders (e.g.
enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
Parkinson's disease), wound repair, cardiovascular disorders (e.g.
endometrial bleeding angio genesis, ischaemias such as coronary
ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,
rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
diabetes and retinal disorders such as retinitis pigmentosum.
The PRO nucleic acids have applications in molecular biology, including
use as hybridization probes, and in chromosome and gene mapping.

50	Sequence	1621 BP; 407 A; 394 C; 374 G; 442 T; 4 other;
	Query Match	87.4%; Score 1359; DB 22; Length 1621;
	Best Local Similarity	92.6%; Pred. No. 0;
	Matches 1494; Conservative	0; Mismatches 4; Indels 115; Gaps 2;
QY	57	AGATGGGGTCTTGGGACCTCTAATTTGCTCTCGTGTATTCGGTGC CGGACTTTCACGAT 116
DB	1	AGATGGCGGTCTTGGGACCTCTAATTTGCTCTCGTGTATTCGGTGC CGGACTTTCACGAT 60
QY	117	GGCTCGCCCAACCTTACTACTCTTCTGTCGCGCTCTCTCTCTGCTGCTTCTCTACTCTGTGA 176
DB	61	GGCTCGCCCAACCTTACTACTCTTCTGTCGCGCTCTCTCTCTGCTGCTTCTCTACTCTGTGA 120
QY	177	GGAACCTGCCCGCGCTCTGCCACGGTCTGCCACCCCAACGAGCGGTAACCCGTTG 236
DB	121	GGAACCTGCCCGCGCTCTGCCACGGTCTGCCACCCCAACGAGCGGTAACCCGTTG 180
QY	237	ACTTTGACTGGAGAGATGGAGATCCTGATGTTTCTTCAGTGCCATTTGATGATGAAGA 296
DB	181	ACTTTGACTGGAGAGATGGAGATCCTGATGTTTCTTCAGTGCCATTTGATGATGAAGA 240
QY	297	ACCGCAGATCCA----- 308
DB	241	ACCGCAGATCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGG 300
QY	309	----- 308
DB	301	CAAACAATTCCTTTCTTCGGCTTGATATTCGCATGGCGCTACTTTACATCACACTCT 360
QY	309	-----TGTTCCTGATGACGTGCAACCCCGCCTATATATGTTGGCCCTGAGTATATCAAGT 362
DB	361	GCATAGTGTTCCTGATGACGTGCAACCCCGCCTATATATGTTGGCCCTGAGTATATCAAGT 420
QY	363	ACTTCAATGATAAAACCATTTGATGAGAACTAGAACGGGACAAGAGGGTCACTTGGATTG 422
DB	421	ACTTCAATGATAAAACCATTTGATGAGAACTAGAACGGGACAAGAGGGTCACTTGGATTG 480
QY	423	TGGAGTTCTTTGGCAATTTGGTCTTAATGACTGCCAATCATTTGGCCCTATCTATGCTGACC 482
DB	481	TGGAGTTCTTTGGCAATTTGGTCTTAATGACTGCCAATCATTTGGCCCTATCTATGCTGACC 540
QY	483	TCTCCCTTAAATACAACCTGTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATA 542
DB	541	TCTCCCTTAAATACAACCTGTACAGGGCTAAATTTTGGGAAGGTGGATGTTTGGACGCTATA 600
QY	543	CTGATGTTACTACGGGTACAAAGTAGACACATCACCCCTCACCAAGCAACTCCCTACCC 602
DB	601	CTGATGTTACTACGGGTACAAAGTAGACACATCACCCCTCACCAAGCAACTCCCTACCC 660
QY	603	TGATCTCTTTCCAAGTGGCAAGGAGCAATGGCGGCCACAGATTGACAAGAAAGGAC 662
DB	661	TGATCTCTTTCCAAGTGGCAAGGAGCAATGGCGGCCACAGATTGACAAGAAAGGAC 720
QY	663	GGGCTGTCTCAT -GGACCTTCTCTGAGGAGAATGTGATCCGAGAAATTAACCTAAATGAG 721
DB	721	GGGCTGTCTCATGGACCTTCTCTGAGGAGAAATGTGATCCGAGAAATTAACCTAAATGAG 780
QY	722	CTATACAGCGGGCCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCCT 781
DB	781	CTATACAGCGGGCCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCCT 840
QY	782	GTGCTTTCAACCCCAACAGTGTCAAGATGGGAAACAAAGAGGATAAATAGATCCCT 841
DB	841	GTGCTTTCAACCCCAACAGTGTCAAGATGGGAAACAAAGAGGATAAATAGATCCCT 900
QY	842	CACCTTTGGCAGTGTCTCTCTCTGTCAATTCAGGCTCTTTCCATTAACCAACGCTGA 901
DB	901	CACCTTTGGCAGTGTCTCTCTCTGTCAATTCAGGCTCTTTCCATTAACCAACGCTGA 960
QY	902	GGCTGCAGCCTTTTATTTATGTTTTTCCCTTTGGCTGTGACTGGTGGGGCAGCATGCAGC 961
DB	961	GGCTGCAGCCTTTTATTTATGTTTTTCCCTTTGGCTGTGACTGGTGGGGCAGCATGCAGC 1020

CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2297 BP; 585 A; 510 C; 546 G; 656 T; 0 other;

Query Match 55.8%; Score 868; DB 22; Length 2297;
Best Local Similarity 100.0%; Pred. No. 2.8e-272;
Matches 868; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

687 AGGAGATGTGATCCGAGAAATTTAACTTAATGAGCTATACAGGGGCCCAAGAACTAT 746
|||||
1419 AGGAGATGTGATCCGAGAAATTTAACTTAATGAGCTATACAGGGGCCCAAGAACTAT 1478
|||||
747 CAAGAGCTGGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACACAGTGT 806
|||||
1479 CAAGAGCTGGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACACAGTGT 1538
|||||
807 CAGATGGGGAAACAAGAGGATAAATAAGATCTCACTTTGGCAGTCTTCTCTCTG 866
|||||
1539 CAGATGGGGAAACAAGAGGATAAATAAGATCTCACTTTGGCAGTCTTCTCTCTG 1598
|||||
867 TCAATTCAGGCTCTTCCCAATACCAAGAGCTGAGGCTTGTATTTATGTTT 926
|||||
1599 TCAATTCAGGCTCTTCCCAATACCAAGAGCTGAGGCTTGTATTTATGTTT 1658
|||||
927 CCCTTTGGCTGTACTGGGTTGGGAGCAGTCAGCTTCTGATTTTAAAGAGCACTAGG 986
|||||
1659 CCCTTTGGCTGTACTGGGTTGGGAGCAGTCAGCTTCTGATTTTAAAGAGCACTAGG 1718
|||||
987 GAATTCAGGACCCCTACAGGAGGCTGCCATGCTGTGGCCAACTGTTTCACTGGAGC 1046
|||||
1719 GAATTCAGGACCCCTACAGGAGGCTGCCATGCTGTGGCCAACTGTTTCACTGGAGC 1778
|||||
1047 AAGAAAGAGATCTCATAGGAGGAGGGGAATGGTTTCCCTCCAAGCTTGGGTCAGTGT 1106
|||||
1779 AAGAAAGAGATCTCATAGGAGGAGGGGAATGGTTTCCCTCCAAGCTTGGGTCAGTGT 1838
|||||
1107 GTTAACCTCTTATCAGCTATTCAGACATCTCCATGGTTTCCCATGAACCTCTGTGGTTT 1166
|||||
1839 GTTAACCTCTTATCAGCTATTCAGACATCTCCATGGTTTCCCATGAACCTCTGTGGTTT 1898
|||||
1167 CATCATCTCTTCTAGTTAGCTTCCAGCTGAGCTTGTAGACCTAGATTTAAACCTTAAGTA 1226
|||||
1899 CATCATCTCTTCTAGTTAGCTTCCAGCTGAGCTTGTAGACCTAGATTTAAACCTTAAGTA 1958
|||||
1227 AGATGCTGGGTATAGAACGCTAAGAAATTTCCCCCAAGGACTCTGTCTTCTTAAAGCCC 1286
|||||
1959 AGATGCTGGGTATAGAACGCTAAGAAATTTCCCCCAAGGACTCTGTCTTCTTAAAGCCC 2018
|||||
1287 TCTGGCTTCTGTTTATGGTCTTCAATTTAAAGTATAAGCCTTAACCTTGTGCTAGTCTTAA 1346
|||||
2019 TCTGGCTTCTGTTTATGGTCTTCAATTTAAAGTATAAGCCTTAACCTTGTGCTAGTCTTAA 2078
|||||
1347 GGAGAAACCTTTACCAACAAGTTTATCATTTGAAGACATATTTGAACACCCCTTAT 1406
|||||
2079 GGAGAAACCTTTACCAACAAGTTTATCATTTGAAGACATATTTGAACACCCCTTAT 2138
|||||
1407 TTGTGGGATTTGAGAGGGGTGAATAGAGGCTTTCAGACTTTTCTTGTGTAGGACTT 1466
|||||

Db 2139 TTGTGGGATTGAGAAGGGTGAATAGAGGCTTGAGACTTTCCTTTTGTGTAGGACTT 2198
Qy 1467 GGAGAGAAATCCCTCGACTTTCCTTAACCTCTGACATACCTCCACACACCACTTGAT 1526
|||||
Db 2199 GGAGAGAAATCCCTCGACTTTCCTTAACCTCTGACATACCTCCACACACCACTTGAT 2258
Qy 1527 GCCTTTCCGTAATAAAAAAGATTGGGATT 1554
|||||
Db 2259 GCCTTTCCGTAATAAAAAAGATTGGGATT 2286
|||||

RESULT 12
ABAL5532
ID ABAL5532 standard; DNA; 3369 BP.
XX
AC ABA15532;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 7863.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.

```
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system
XX cancers and metastases -
XX
XX Disclosure; SEQ ID NO 7863; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-AB21534) and proteins
XX (AB14678-AB18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemia;
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3369 BP; 850 A; 724 C; 788 G; 1007 T; 0 other;
XX
XX Query Match 55.8%; Score 868; DB 22; Length 3369;
XX Best Local Similarity 100.0%; Pred. No. 3.5e-272;
XX Matches 868; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 687 AGGAGATGTGATCCGAGAAATTAACCTTAATGAGCTATACCAGGGGCCAAGAACTAT 746
Db 2491 AGGAGATGTGATCCGAGAAATTAACCTTAATGAGCTATACCAGGGGCCAAGAACTAT 2550
QY 747 CAAAGCTGGAGACAATATCCCTGAGGACGCTGTGGCTTCAACCCACCACAGTGT 806
Db 2551 CAAAGCTGGAGACAATATCCCTGAGGACGCTGTGGCTTCAACCCACCACAGTGT 2610
```


XX WIPI; 2001-541565/60.

XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -

XX Disclosure; SEQ ID NO 7864; 1701pp + Sequence Listing; English.

XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins
CC (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX CC
SQ Sequence 3369 BP; 849 A; 724 C; 789 G; 1007 T; 0 other;

Query Match 55.8%; Score 868; DB 22; Length 3369;
Best Local Similarity 100.0%; Pred. No. 3.5e-272;
Matches 868; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 687 AGGAGATCTGTCCCGAGAATTAACTTAACCTAATCAGCTATACCAGCGGGCCCAAGAACTAT 746
DB 2491 AGGAGATCTGTCCCGAGAATTAACTTAACCTAATCAGCTATACCAGCGGGCCCAAGAACTAT 2550
QY 747 CAAAAGCTGGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCCCACACACTGT 806
DB 2551 CAAAAGCTGGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCCCACACACTGT 2610
QY 807 CAGATGGGGAAAAACAAGAGGATAAATAAGATCCTCCTTTGGCAGTGCTTCCTCCTCTG 866
DB 2611 CAGATGGGGAAAAACAAGAGGATAAATAAGATCCTCCTTTGGCAGTGCTTCCTCCTCTG 2670
QY 867 TCAATTCCAGGCTCTTTTCCATAACCAACAGCCCTGAGGCTGCAGCCCTTTATTTATGTTTT 926
DB 2671 TCAATTCCAGGCTCTTTTCCATAACCAACAGCCCTGAGGCTGCAGCCCTTTATTTATGTTTT 2730
QY 927 CCTTTGGCTGTGACTGGGTGGGCGAGCATGCGAGCTTCGATTTTAAAGAGGCACTTAGG 986
DB 2731 CCTTTGGCTGTGACTGGGTGGGCGAGCATGCGAGCTTCGATTTTAAAGAGGCACTTAGG 2790
QY 987 GAATTGTCAGGCCACCCCTACAGGAAGCCCTGCCATGCTGTGGCCAACCTGTTTCACTGGAGC 1046
DB 2791 GAATTGTCAGGCCACCCCTACAGGAAGCCCTGCCATGCTGTGGCCAACCTGTTTCACTGGAGC 2850
QY 1047 AAGAAAGAGATCTCATAGACGAGGAGGGGNAATGTTTCCCTCCAGCTGGGTCAGTGT 1106
DB 2851 AAGAAAGAGATCTCATAGACGAGGAGGGGNAATGTTTCCCTCCAGCTGGGTCAGTGT 2910
QY 1107 GTTAACCTGCTTATCAGCTATTTCAGACATCTCCATGTTTCTCCATGAACCTCTGTGGTTT 1166
DB 2911 GTTAACCTGCTTATCAGCTATTTCAGACATCTCCATGTTTCTCCATGAACCTCTGTGGTTT 2970
QY 1167 CATCATTTCTTCTTAGTTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCCTAAGGTA 1226
DB 2971 CATCATTTCTTCTTAGTTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCCTAAGGTA 3030
QY 1227 AGATGCTGGGGTATAGAACGCTAAGAAATTTTCCCACCAAGACGCTCTGTCTCTTAAGCCC 1286
DB 3031 AGATGCTGGGGTATAGAACGCTAAGAAATTTTCCCACCAAGACGCTCTGTCTCTTAAGCCC 3090
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI

QY 1287 TTCTGGCTTCGTTTATGCTCTTCATTAATAAGTATTAAGCCTACTTGTGCTAGTCCCTAA 1346
 DB 3091 TTCTGGCTTCGTTTATGCTCTTCATTAATAAGTATTAAGCCTACTTGTGCTAGTCCCTAA 3150
 QY 1347 GGAGAAACCTTTAACCCACAAAGTTTATCATTTGAAGACAATATTGAACAACCCCTATT 1406
 DB 3151 GGAGAAACCTTTAACCCACAAAGTTTATCATTTGAAGACAATATTGAACAACCCCTATT 3210
 QY 1407 TTGTGGGATGAGAGGGGCTAATAGAGGCTTGAGACTTCTCTTTGTGTAGACTTT 1466
 DB 3211 TTGTGGGATGAGAGGGGCTAATAGAGGCTTGAGACTTCTCTTTGTGTAGACTTT 3270
 QY 1467 GGAGAGAAATCCCTGACCTTCTACTAACCTCTGACATACCTCCACACCCAGTTGAT 1526
 DB 3271 GGAGAGAAATCCCTGACCTTCTACTAACCTCTGACATACCTCCACACCCAGTTGAT 3330
 QY 1527 GGCCTTCGTAATAAAAGATTGGGATT 1554
 DB 3331 GGCCTTCGTAATAAAAGATTGGGATT 3358

RESULT 14
 \15981

AAAL5981 standard; cDNA; 774 BP.

XX AC AAAL5981;
 XX DT

XX 12-JUN-2000 (first entry)
 XX DE

XX Human protein clone HP10392 coding sequence.
 XX KW

Human protein; hydrophobic domain; nutritional source; haematopoiesis;
 cytokine production; cell proliferation; cell differentiation;
 immune deficiency; infectious disease; autoimmune disorder; asthma;
 multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
 allergic reaction; osteoporosis; osteoarthritis; periodontal disease;
 nervous system disorder; Alzheimer's disease; Parkinson's disease;
 Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;
 systemic cytokine damage; tissue differentiation; contraceptive condition;
 coagulation disorder; myocardial infarction; inflammatory condition; stroke;
 septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;
 nephritis; therapy; ss.

XX OS Homo sapiens.
 XX PN

XX WO200005367-A2.
 XX PD

XX 03-FEB-2000.
 XX PF

XX 22-JUL-1999; 99WO-JP03929.
 XX PR

XX 24-JUL-1998; 98JP-0208820.
 XX PR

XX 07-AUG-1998; 98JP-0224105.
 XX PR

XX 25-AUG-1998; 98JP-0238116.
 XX PR

XX 09-SEP-1998; 98JP-0254736.
 XX PR

XX 29-SEP-1998; 98JP-0275505.
 XX PA

XX (SAGA) SAGAMI CHEM RES CENT.
 XX PA (PROT-) PROTEGENE INC.

XX Kato S, Kimura T;
 XX PI

XX WPI; 2000-182694/16.
 XX DR

XX P-PSDB; AAY94893.
 XX PT

XX Novel human proteins having hydrophobic domains useful for treating
 osteoporosis, Alzheimer's disease, Parkinson's disease, asthma,
 multiple sclerosis, rheumatoid arthritis, cancer, anaemia, and stroke

XX Claim 3; Page 328; 35lpp; English.
 XX PS

XX This sequence encodes a human protein of the invention, which has
 XX CC

CC hydrophobic domains. The DNA sequences can be used as a probe or as a
 CC genetic marker. The protein can also be used as a marker, and to identify
 CC potential genetic disorders. The DNA and protein can also be used as
 CC nutritional sources or supplements. The protein exhibits cytokine, cell
 CC proliferation, cell differentiation activities and induces production of
 CC other cytokines in certain cell populations. The protein also exhibits
 CC immune stimulating or immune suppressing activity. It can be used in the
 CC treatment of various immune deficiencies and disorders, and to treat
 CC infectious diseases caused by viral, bacterial, fungal or other
 CC infections. The protein is also used for treating autoimmune disorders
 CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid
 CC arthritis. It is also useful in the treatment of allergic reactions and
 CC conditions such as asthma, and in immune suppression after organ
 CC transplantation. The protein is useful in regulation of haematopoiesis
 CC and consequently in the treatment of myeloid or lymphoid cell
 CC deficiencies. It is also used in compositions for tissue growth or
 CC regeneration. The protein is also used in the treatment of osteoporosis
 CC or osteoarthritis and in the treatment of periodontal disease and other
 CC tooth repair processes. The protein is used in the treatment of nervous
 CC system disorders such as Alzheimer's disease, Parkinson's disease, and
 CC Huntington's disease. They are useful for protection or regeneration and
 CC treatment of lung or liver fibrosis, reperfusion injury in various
 CC tissues, and conditions resulting from systemic cytokine damage. They are
 CC also used for promoting or inhibiting tissue differentiation. They are
 CC also used as contraceptives since they exhibit activin or inhibin related
 CC activities and as a fertility inducing therapeutic. They are used for
 CC treating various coagulation disorders and in treatment and prevention of
 CC conditions resulting from coagulation activities e.g. myocardial
 CC infarction or stroke. They also acts as receptors, receptor ligands or
 CC inhibitors or agonists of receptor/ligand interactions. They are used to
 CC treat inflammatory conditions such as septic shock, sepsis, ischaemia
 CC reperfusion injury, arthritis, and nephritis. They can be used to
 CC prevent tumours.

XX SQ Sequence 774 BP; 200 A; 200 C; 190 G; 184 T; 0 other;

Query Match 49.8%; Score 774; DB 21; Length 774;
 Best Local Similarity 100.0%; Pred. No. 7.5e-242;
 Matches 774; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 59 ATGCGGGTCTGGCACCTCTAATGCTCTCGTGTATTCGGTGCCGGACATTTCCACGATGG 118
 DB 1 ATGCGGGTCTGGCACCTCTAATGCTCTCGTGTATTCGGTGCCGGACATTTCCACGATGG 60
 QY 119 CTCGCCCCACCTTACTACTCTTCTGCGCCCTCTCTCTGCTGCTTCTTCTACTCTGAGG 178
 DB 61 CTCGCCCCACCTTACTACTCTTCTGCGCCCTCTCTCTGCTGCTTCTTCTACTCTGAGG 120
 QY 179 AAATCTGCGCGCGCTCTGCGCACGGTCTGCCACCCACGGAAGAGCGGTAAACCGGTGAC 238
 DB 121 AAATCTGCGCGCGCTCTGCGCACGGTCTGCCACCCACGGAAGAGCGGTAAACCGGTGAC 180
 QY 239 TTTGACTGGAGAGAGTGGAGATCCTGATGTTTCTCAGTGCCTATTTGTATGATGAAGAAC 298
 DB 181 TTTGACTGGAGAGAGTGGAGATCCTGATGTTTCTCAGTGCCTATTTGTATGATGAAGAAC 240
 QY 299 CGCAGATCCATGTTCTCTGATGACGTGCAAAACCCCTATATATGCGCCCTCAGTATATC 358
 DB 241 CGCAGATCCATGTTCTCTGATGACGTGCAAAACCCCTATATATGCGCCCTCAGTATATC 300
 QY 359 AAGTACTTCAATGATATAAAACCAATTGATGAGGAAGTACGAGGCAAGAGGCTCACTTGG 418
 DB 301 AAGTACTTCAATGATATAAAACCAATTGATGAGGAAGTACGAGGCAAGAGGCTCACTTGG 360
 QY 419 ATTTGGGAGTTCTTTTGCCCAATTTGGTCTAATGACTGCCAATCATTTGCCCTTCTATGCT 478
 DB 361 ATTTGGGAGTTCTTTTGCCCAATTTGGTCTAATGACTGCCAATCATTTGCCCTTCTATGCT 420
 QY 479 GACCTCTCCCTTAATACAACTGTACAGGCTAAATTTTGGGAAGTGTGATTTGGACGC 538
 DB 421 GACCTCTCCCTTAATACAACTGTACAGGCTAAATTTTGGGAAGTGTGATTTGGACGC 480
 QY 539 TATACTGATGTTAGTACGGGTACAAAGTACGACATCACCCCTCACCAGCACTCCCT 598

Db 481 TATACTGATGTAGTACCGGTACAAAGTGAGCACATCACCCCTCACCAAGCAACTCCCT 540
Qy 599 ACCCTGATCCCTGTTCCCAAGGTGGCAAGGAGCAATGCGGCGCCACAGATTGACAAGAAA 558
Db 541 ACCCTGATCCCTGTTCCCAAGGTGGCAAGGAGCAATGCGGCGCCACAGATTGACAAGAAA 600
Qy 659 GGACGGGCTGCTCATGACCTTCTCTGAGGAGAAATGTATCCGAGAAATTTAACTTAAAT 718
Db 601 GGACGGGCTGCTCATGACCTTCTCTGAGGAGAAATGTATCCGAGAAATTTAACTTAAAT 560
Qy 719 GAGCTATACCAAGCGGCGCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAG 778
Db 661 GAGCTATACCAAGCGGCGCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAG 720
Qy 779 CCTGTGGCTTCAACCCCAACACAGTGTCTGAGTGGGGAACAAGAGATAAA 832
721 CCTGTGGCTTCAACCCCAACACAGTGTCTGAGTGGGGAACAAGAGATAAA 774

RESULT 15

AAS83076
ID AAS83076 standard; cDNA; 2030 BP.

XX AAS83076;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #18880.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

X Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR P-PSDB: ABG18889.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX Claim 1; SEQ ID No 18880; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2030 BP; 500 A; 476 C; 516 G; 523 T; 15 other;

Qy Query Match 39.4%; Score 613; DB 23; Length 2030;
Db Best Local Similarity 78.7%; Pred. No. 5.5e-189;
Qy Matches 1403; Conservative 0; Mismatches 150; Indels 230; Gaps 47;

Qy 2 GGGGAGGCGGGCGGAGACCTACGACCGCGGAGAGTGGCCCTTACGGCCGAAAGATG 61
Db 44 GGGGAGGCGGGCGGAGACCTACGACCGCGGAGAGTGGCCCTTACGGCCGAAAGATG 103
Qy 62 GGGGCTTTTGGCACCTCTAATT--GCTCTCTGTGTATTTCGGTGGCGCGAC--TTTCACGATGG 118
Db 104 GCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTTCGGTGGCGCGACATTTTCACGATGG 163
Qy 119 CTCGCCCAACCTTACTACTCTCTGTCGG--CCCTGCTCTCTGTCGCTT--CCTACTCTGTA 176
Db 164 CTCGCCCAACCTTACTACTCTCTGTCGGCCCTCTCTCTGTCGCTTCTCCTACTCTGTA 223
Qy 177 GGAACCTGCCGCGCTCTGCCAGGTCTGCCACCCAA--CGCGAAGACGGTA--CCCGTG 234
Db 224 GGAACCTGCCGCGCTCTGCCAGGTCTGCCACCCAA--CGCGAAGACGGTA--CCCGTG 283
Qy 235 TGACTTTGACTGGAGAG--AAGTGGAGATCTCTGATGTTTCTCAGTGCATTGTGATGTA 293
Db 284 TGACTTTGACTGGAGAGAGTGGAGATCTCTGATGTTTCTCAGTGCATTGTGATGTA 343
Qy 294 AGAA--CCGCAGATCCA----- 308
Db 344 AGAACCAGATGATCCATCTCTGGAGCAACATATAGCAACATTTTTCATGTTTAAAGTA 403
Qy 309 ----- 308
Db 404 AAGTGGCCAAACAATTTCTTTCTCCCGCTTGGATATTGCGATGGCGCTACTTTACATC 463
Qy 309 -----TGTTCTCTGATGAGTGCACCAACCC-----CCCTATATATGCGCCTG 350
Db 464 ACACCTCTGATAGTGTCTCTGATGAGTGCACCAACCCCGCTCTCCCATATATGCGCCTG 523
Qy 351 AGTATATCAAGTACTTCAAT--GATAAACCATTGATGAGAACTAGAACGGGACAGAG 409
Db 524 AGTATATCAAGTACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 583
Qy 410 GTCACCTGGA--TTGTGGAGTCTTTGCCAATTTGGTCTTAATGACTGCCAATCAATTTGCC 467
Db 584 GTCAATTTGGAATTTGTTGGAGTCTTTTGGCAATTTGTTGTAATGACTGGCAATCAATTTGCC 643
Qy 468 CTATCTATCTGACCTCTCCCTTAAATACAACTGTACAGGCTTAAATTTTGGGAAGTGG 527
Db 644 CTATCTATCTGACCTCTCCCTTAAATACAACTGTACAGGCTTAAATTTTGGGAAGTGG 703
Qy 528 ATGTTGGAGCTATCTGATGTTAGTACGCGGTACAAAGTGAAGTGAACATCACCCCTCACCA 587
Db 704 ATGTTGGAGCTATCTGATGTTAGTACGCGGTACAAAGTGAAGTGAACATCACCCCTCACCA 763
Qy 588 A--GCAACTCCCTTACCTT--GATCCTGTTTCAAGTGGCAA--GGAGGCAATGCGGCGGCC 642
Db 764 AGGCAACTCCCTTACCTTACCTTCCCAAGTGGGCAAGGAGGCAATGCGGCGGCC 823
Qy 643 ACAGATT-----GACAAGAAAGGCGGCTGTCTCATGGACCTTCTCTG----- 686
Db 824 ACAGATTGACAAGAAAGGCGGCTGTTTTCATGGGACCTTCTCTGAGTACCTGA 883
Qy 687 -----AGGAGAATGTATCCGAGAAATTAACCTAAATGAGCTATACCAAGCGGCC 736
Db 884 AAGGAAGGCGAGAGATGTATCCGAGAAATTAACCTAAATGAGCTATACCAAGCGGCC 943

coagulation disorder; myocardial infarction; inflammatory condition;
septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;
nephritis; therapy; ss.

Homo sapiens.

WO200005367-A2.

03-FEB-2000.

22-JUL-1999; 99WO-JP03929.

24-JUL-1998; 98JP-0208820.

07-AUG-1998; 98JP-0224105.

25-AUG-1998; 98JP-0238116.

09-SEP-1998; 98JP-0254736.

29-SEP-1998; 98JP-0275505.

(SAGA) SAGAMI CHEM RES CENT.

(PROT-) PROTEGENE INC.

Kato S, Kimura T;

WPI; 2000-182694/16.

P-PSDB; AAY94893.

Novel human proteins having hydrophobic domains useful for treating
osteoporosis, Alzheimer's disease, Parkinson's disease, asthma,
multiple sclerosis, rheumatoid arthritis, cancer, anaemia, and stroke -

Claim 3; Page 328; 351pp; English.

This sequence encodes a human protein of the invention, which has
hydrophobic domains. The DNA sequences can be used as a probe or as a
genetic marker. The protein can also be used as a marker, and to identify
potential genetic disorders. The DNA and protein can also be used as
nutritional sources or supplements. The protein exhibits cytokine, cell
proliferation, cell differentiation activities and induces production of
other cytokines in certain cell populations. The protein also exhibits
immune stimulating or immune suppressing activity. It can be used in the
treatment of various immune deficiencies and disorders, and to treat
infectious diseases caused by viral, bacterial, fungal or other

infections. The protein is also used for treating autoimmune disorders
such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid
arthritis. It is also useful in the treatment of allergic reactions and
conditions such as asthma, and in immune suppression after organ
transplantation. The protein is useful in regulation of haematopoiesis
and consequently in the treatment of myeloid or lymphoid cell

deficiencies. It is also used in compositions for tissue growth or
regeneration. The protein is also used in the treatment of osteoporosis
or osteoarthritis and in the treatment of periodontal disease and other
tooth repair processes. The protein is used in the treatment of nervous
system disorders such as Alzheimer's disease, Parkinson's disease, and
Huntington's disease. They are useful for protection or regeneration and
treatment of lung or liver fibrosis, reperfusion injury in various
tissues, and conditions resulting from systemic cytokine damage. They are
also used for promoting or inhibiting tissue differentiation. They are
also used as contraceptives since they exhibit activin or inhibin related
activities and as a fertility inducing therapeutic. They are used for
treating various coagulation disorders and in treatment and prevention of
conditions resulting from coagulation activities e.g. myocardial
infarction or stroke. They also acts as receptors, receptor ligands or
inhibitors or agonists of receptor/ligand interactions. They are used to
treat inflammatory conditions such as septic shock, sepsis, ischaemia
reperfusion injury, arthritis, and nephritis. They can be used to
prevent tumours.

Sequence 774 BP; 200 A; 200 C; 190 G; 184 T; 0 other;

Alignment Scores:
Pred. No.: 3,41e-155 Length: 774

Score: 1359.00 Matches: 258

Percent Similarity: 100.00 Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-954-846-2 (1-258) x AAA15981 (1-774)

QY 1 MetAlaValLeuAlaProLeuLeuAlaLeuValTyrSerValProArgLeuSerArgTTP 20
DB 1 ATGGCGGCTTGGCACCCTCTAATGTCTCGTGATTCGGTGGCGGACCTTTCACGATGG 60
QY 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArg 40
DB 61 CTGCGCCAAACCTTACTACCTTCTGTGCGCCCTCTCTCTGCTGCTCTTCTACTCTGAGG 120
QY 41 LysLeuProLeuLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
DB 121 AAACCTGCCCGCGCTCTGCCACGCTCTGCCACCAACGCGGAGACCGGTACCGGTGAC 180
QY 61 PheAspTyrArgGluValGluLeuMetPheLeuSerAlaAlaValMetMetLysAsp 80
DB 181 TTTGACTGGAGAGAGAGTGGAGATCCGTCATGTTTCTCAGTGCCCATTTGTGATGATGAAGAAC 240
QY 81 ArgArgSerMetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIle 100
DB 241 CGCAGATCCATGTCTCTGATGACGTGCAAAACCCCTATATATGGCCCTGAGTATATC 300
QY 101 LysTyrPheAsnAspLysThrIleAspGluLeuGluArgAspLysArgValThrTrp 120
DB 301 AAGTACTTCAATGATAAAAACCATTTGATGAGGAAGTAAAGACGGGACGAGGGTCACTTGG 360
QY 121 IleValGluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAla 140
DB 361 ATGTGGAGTCTTTCGCCAATTTGGTCTAATGACTGCCAATCAATTCGCCCTATCTATGCT 420
QY 141 AspLeuSerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArg 160
DB 421 GACCTCTCCCTTAATACAACTGTACAGGCTAAATTTGGGAAGGTGGATGTTGGACGC 480
QY 161 TyrThrAspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuPro 180
DB 481 TATACTGATGTTAGTACGGGTACAAAGTGACACATCACCCCTCACCAAGCACTCCCT 540
QY 181 ThrLeuLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLys 200
DB 541 ACCCTGATCCTGTTCAGAGGTGGCAAGGAGGAATCGCGCGGCACAGATTGACAAAGAA 600
QY 201 GlyArgAlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsn 220
DB 601 GGACGGGCTGTCTCATGGACCTTCTCTGAGGAGATGTGATCGGAGAAATTAACCTAAT 660
QY 221 GluLeuTyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGln 240
DB 661 GAGCTATACAGCGGCGCCAAAGAACTATCAAAAGCTGGAGACAATATCCCTGAGGAGCAG 720
QY 241 ProValAlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
DB 721 CCTGTGGCTTCAACCCCGCCACACAGTGTGAGATGGGGGAAACAAAGAGGATAAA 774

RESULT 2

AAA15991

ID AAA15991 standard; cDNA; 1527 BP.

XX AAA15991;

AC AAA15991;

XX 12-JUN-2000 (first entry)

XX Human protein clone HP10392 full length coding sequence.

XX Human protein; hydrophobic domain; nutritional source; haematopoiesis;

KW cytokine production; cell proliferation; cell differentiation;

KW immune deficiency; infectious disease; autoimmune disorder; asthma;

KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;

KW allergic reaction; osteoporosis; osteoarthritis; periodontal disease;

nervous system disorder; Alzheimer's disease; Parkinson's disease; Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury; systemic cytokine damage; tissue differentiation; contraceptive; stroke; coagulation disorder; myocardial infarction; inflammatory condition; septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour; nephritis; therapy; ss.

Homo sapiens.

WO200005367-A2.

03-FEB-2000.

22-JUL-1999; 99WO-JP03929.

24-JUL-1998; 98JP-0208920.

07-AUG-1998; 98JP-0224105.

25-AUG-1998; 98JP-0238116.

09-SEP-1998; 98JP-0254736.

29-SEP-1998; 98JP-0275505.

(SAGA) SAGAMI CHEM RES CENT.
(PROT-) PROTEGENE INC.

Kato S, Kimura T;

WPI; 2000-182694/16.

P-PSDB; AAY94893.

Novel human proteins having hydrophobic domains useful for treating osteoporosis, Alzheimer's disease, Parkinson's disease, asthma, multiple sclerosis, rheumatoid arthritis, cancer, anaemia, and stroke - Claim 4; Page 341-343; 351pp; English.

This sequence encodes a human protein of the invention, which has hydrophobic domains. The DNA sequences can be used as a probe or as a genetic marker. The protein can also be used as a marker, and to identify potential genetic disorders. The DNA and protein can also be used as nutritional sources or supplements. The protein exhibits cytokine, cell proliferation, cell differentiation activities and induces production of other cytokines in certain cell populations. The protein also exhibits immune stimulating or immune suppressing activity. It can be used in the treatment of various immune deficiencies and disorders, and to treat infectious diseases caused by viral, bacterial, fungal or other infections. The protein is also used for treating autoimmune disorders such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid arthritis. It is also useful in the treatment of allergic reactions and conditions such as asthma, and in immune suppression after organ transplantation. The protein is useful in regulation of haematopoiesis and consequently in the treatment of myeloid or lymphoid cell deficiencies. It is also used in compositions for tissue growth or regeneration. The protein is also used in the treatment of osteoporosis or osteoarthritis and in the treatment of periodontal disease and other tooth repair processes. The protein is used in the treatment of nervous system disorders such as Alzheimer's disease, Parkinson's disease, and Huntington's disease. They are useful for protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage. They are also used for promoting or inhibiting tissue differentiation. They are also used as contraceptives since they exhibit activin or inhibin related activities and as a fertility inducing therapeutic. They are used for treating various coagulation disorders and in treatment and prevention of conditions resulting from coagulation activities e.g. myocardial infarction or stroke. They also acts as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. They are used to treat inflammatory conditions such as septic shock, sepsis, ischaemia reperfusion injury, arthritis, and nephritis. They can be used to prevent tumours.

Sequence 1527 BP; 382 A; 373 C; 361 G; 411 T; 0 other;

Alignment Scores:

Pred. No.: 9.32e-155
Score: 1359.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 21

Length: 1527
Matches: 258
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-954-846-2 (1-258) x AAA15991 (1-1527)

Qy 1 MetaValLeuAlaProLeuLeuAlaLeuValTyrSerValProArgLeuSerArgTrp 20
Db 25 ATGGCGCTCTGGCACCTCTAAATGCTCTCGTGTATTGGTGGCGGAGATTCACGATGG 84
Qy 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArg 40
Db 85 CTCGCCCAACCTTACTACCTTCTCGGCCCTGCTCTGCTGCTTCTCTACTCTGAGG 144
Qy 41 LysLeuProLeuProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
Db 145 AAACCTGGCGGCTCTGCCACGGTCTGCCACCCCAACCGGAGAGCGTAACCCGCTGAC 204
Qy 61 PheAspTrpArgGluValGluLeuLeuMetPheLeuSerAlaLeuValMetMetLysAsn 80
Db 205 TTTGACTGGAGAGAAGTGGAGATCTCTGATGTTTCTCACTGATGATGATGATGATGAT 264
Qy 81 ArgArgSerMetPheLeuMetThrCysLysProLeuTyrMetGlyProGluTyrIle 100
Db 265 CGCAGATCCATGTTCTCTGATGACGTGCAAAACCCCTATATATGGGCCCTGAGTATATC 324
Qy 101 LysTyrPheAsnAspLysThrIleAspGluGluLeuGluAspLysArgValThrTrp 120
Db 325 AAGTACTTCAATGATATAAACCATTTGATGAGGAAGTAGAACGGGACAGAGGCTACTTGG 384
Qy 121 IleValGluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAla 140
Db 385 ATTGTGGAGTTCTTTGGCAATTTGCTCTAATGACTGCCAATCATTTGCCCTCTATCTGCT 444
Qy 141 AspleuSerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArg 160
Db 445 GACCTCTCCCTTAAATCACTGTACAGGGCTAAATTTTGGGAAGGTGGATTTGGAGCGC 504
Qy 161 TyrThrAspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuPro 180
Db 505 TATACTGATGTTAGTACGGGTACAAAGTAGACACATCACCCCTCACCAACCACTCCCT 564
Qy 181 ThrLeuIleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLys 200
Db 565 ACCCTGATCCTGTTCCAAAGGTGGCAAGGAGCAATGGCGGCCGCACAGATTGACAAGAA 624
Qy 201 GlyArgAlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsn 220
Db 625 GGACGGGCTGTCTCATGGACCTTCTCTGAGGAGAATGTGATCCGAGAAATTTAACTTAAT 684
Qy 221 GluLeuTyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGln 240
Db 685 GAGCTATACACGGCGGCCCAAGAACTATCAAAAGGCTGGAGACAATATCCCTGAGGAGAG 744
Qy 241 ProValAlaSerThrProThrValSerAspGlyGluAsnLysLysAspLys 258
Db 745 CCTGTGCTTCAACCCCCACACAGTGTGATGGGGAACAAACAGAGAGATAAA 798

RESULT 3

AAAF16316

ID AA16316 standard; cDNA: 1598 BP.

AC AA16316;

XX 13-MAR-2001 (first entry)

DT 13-MAR-2001 (first entry)

XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:751.

DE Human prostate cancer; prostate cancer antigen; detection; diagnosis;

KW Human prostate cancer; prostate cancer antigen; detection; diagnosis;

KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;

KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.

XX Homo sapiens.
XX WO200055174-A1.
XX PN
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US05988.
XX PR 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.

XX PI Rosen CA, Ruben SM;
XX DR WPI: 2000-587513/55.
XX DR P-PSDB; AAB57113.

Prostate cancer associated gene sequences, referred to as prostate cancer antigens, useful for treatment, prevention, and diagnosis of disorders such as prostate cancer -

XX Claim 1: Page 1172; 2338pp; English.
XX AAF15566 to AAF16505 encode the human prostate cancer associated proteins, called prostate cancer antigens, given in AAB56363 to AAB57302. The prostate cancer antigens can have neuroprotective, cytostatic, cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal, nephrotropic, antiinfective, gynaecological and antibacterial activities, and can be used in gene therapy. The prostate cancer antigen polynucleotides may be used for detection of prostate cancer, chromosome identification, as chromosome markers, and for numerous other diagnostic or research purposes. The prostate cancer antigens may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to AAB57303 represent sequences used in the exemplification of the present invention.

SQ Sequence 1598 BP; 437 A; 381 C; 365 G; 414 T; 1 other;

Alignment Scores:
Pred. No.: 9.96e-155 Length: 1598
Score: 1359.00 Matches: 258
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
ry Match: 100.00% Indels: 0
Gaps: 21

US-09-954-846-2 (1-258) x AAF16316 (1-1598)

QY 1 MetAlaValLeuAlaProLeuLeuAlaLeuValTyrSerValProArgLeuSerArgTrp 20
Db 43 ATGGCGGTCTTGGACCTCTAATGTCTCGTGTATTCGGTGGCGGACTTTCACCATGG 102
QY 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArg 40
Db 103 CTCGCCAACCTTACTACCTCTGTGGCCGCTCTCTGCTGCTTCTCTACTCTGTGAGG 162
QY 41 LysLeuProLeuProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
Db 163 AAATGCCCGCTGTGCGCGTGTGCGCACCAACCGGAGACGTAACCGTGTGAC 222
QY 61 PheAspTrpArgGluValGluLeuLeuMetPheLeuSerAlaIleValMetLysAsp 80
Db 223 TTGTACTGGAGAGAGAGTGGAGATCCCTGATGTTTCTCAGTGCCCATTTGATGATGAAGAAC 282
QY 81 ArgArgSerMetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIle 100

Db 283 CGCAGATCCATGTTCTCTGATGACGTGCAAAACCCCTATATATATGGCCCTGAGTATATC 342
QY 101 LysTyrPheAsnAspLysThrIleAspGluGluLeuGluArgAspLysArgValThrTrp 120
Db 343 AAGTACTTCAATGATAAAACCATTTGATGAGGAACGTAGAACGGGACAGGGTCACTTGG 402
QY 121 IleValGluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAla 140
Db 403 ATTGTGGAGTCTTTGCCAATTTGCTTAATGACTGCCAATCATTTGCCCTTATCTATGCT 462
QY 141 AspLeuSerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValaspValGlyArg 160
Db 463 GACCTCTCCCTTAATAACACTCTACAGGCGCTAAATTTTGGGAAGGTGGATTTGGAGCGC 522
QY 161 TyrThrAspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuPro 180
Db 523 TATACTGATGTTAGTACGGCGGTACAAAGTGAGGACATCACCCCTCACAAGCAACTCCCT 582
QY 181 ThrLeuIleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLys 200
Db 583 ACCCTGATCTGTTCCAAAGTGCAAGGAGGCAATGCGGCGGCACAGATTGACAAGAAA 642
QY 201 GlyArgAlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsn 220
Db 643 GGACGGGCTGCTCTCATGGACCTTCTCTGAGGAGATGTGATCCGAGAAATTTAACTTAAT 702
QY 221 GluLeuTyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGln 240
Db 703 GAGCTATACAGCGGCGCAAGAAATATCAAGGCTGGAGACAATATCCCTGAGGAGCAG 762
QY 241 ProValAlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
Db 763 CCTGTGCTTCAACCCCAACCCACAGTGTGATGGGGGAAACAAAGAGGATAAA 816
RESULT 4
AA52251
ID AAX52251.standard; DNA; 1618 BP.
XX AAX52251;
XX DT 25-JUN-1999 (first entry)
XX DE Protein PRO270 cDNA clone DNA39510-1181.
XX KW Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophla areata;
KW anti-thrombotic; wound healing; tissue repair; ss.
XX OS Homo sapiens.
XX PN WO9914328-A2.
XX PD 25-MAR-1999.
XX PF 16-SEP-1998; 98WO-US19330.
XX PR 25-NOV-1997; 97US-0066840.
XX PR 17-SEP-1997; 97US-0059113.
XX PR 17-SEP-1997; 97US-0059115.
XX PR 17-SEP-1997; 97US-0059117.
XX PR 17-SEP-1997; 97US-0059119.
XX PR 17-SEP-1997; 97US-0059121.
XX PR 17-SEP-1997; 97US-0059122.
XX PR 17-SEP-1997; 97US-0059184.
XX PR 18-SEP-1997; 97US-0059263.
XX PR 15-OCT-1997; 97US-0062125.
XX PR 17-OCT-1997; 97US-0062285.

PR 17-OCT-1997; 97US-0062287.
PR 21-OCT-1997; 97US-0063486.
PR 24-OCT-1997; 97US-0062814.
PR 24-OCT-1997; 97US-0062816.
PR 24-OCT-1997; 97US-0063045.
PR 24-OCT-1997; 97US-0063120.
PR 24-OCT-1997; 97US-0063121.
PR 24-OCT-1997; 97US-0063127.
PR 24-OCT-1997; 97US-0063128.
PR 24-OCT-1997; 97US-0063329.
PR 27-OCT-1997; 97US-0063327.
PR 27-OCT-1997; 97US-0063541.
PR 28-OCT-1997; 97US-0063542.
PR 28-OCT-1997; 97US-0063544.
PR 28-OCT-1997; 97US-0063549.
PR 28-OCT-1997; 97US-0063550.
PR 28-OCT-1997; 97US-0063564.
PR 29-OCT-1997; 97US-0063435.
PR 29-OCT-1997; 97US-0063704.
PR 29-OCT-1997; 97US-0063732.
PR 29-OCT-1997; 97US-0063738.
PR 29-OCT-1997; 97US-0063734.
PR 29-OCT-1997; 97US-0064215.
PR 29-OCT-1997; 97US-0063735.
PR 31-OCT-1997; 97US-0063870.
PR 31-OCT-1997; 97US-0064103.
PR 03-NOV-1997; 97US-0064248.
PR 07-NOV-1997; 97US-0064809.
PR 12-NOV-1997; 97US-0065186.
PR 17-NOV-1997; 97US-0065846.
PR 18-NOV-1997; 97US-0065693.
PR 21-NOV-1997; 97US-0066120.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066772.
PR 24-NOV-1997; 97US-0066466.
PR 24-NOV-1997; 97US-0066770.
PR 24-NOV-1997; 97US-0066511.
PR 24-NOV-1997; 97US-0066453.

(GETH) GENENTECH INC.

Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;

WPI; 1999-229533/19.

P-PSDB; AAY13380.

New isolated human genes and polypeptides used in, e.g. treatment of gastrointestinal ulceration

Claim 2; Fig 75; 320pp; English.

AA52213-74 encode secreted and transmembrane human proteins, and are obtained from cDNA libraries, prepared from fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina. The encoded polypeptides have specific uses based on their homology to known polypeptides, e.g. PRO211 and PRO217 can be used for disorders associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial cancers such as lung squamous cell carcinoma of the vulva and gliomas), potent effects on cell growth and development, diseases related to growth or survival of nerve cells including Parkinson's disease, Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used as a target for anti-tumor drugs. PRO533 may be used in the treatment of Usher Syndrome or Atrophia areata; PRO269 can be used as an anti-thrombotic agent; PRO287 polypeptides and portions may have therapeutic applications in wound healing and tissue repair; PRO317 can be used for treating problems of the kidney, uterus, endometrium, blood vessels, or related tissue, e.g. in the heart of genital tract.

XX

SQ Sequence 1618 BP; 406 A; 394 C; 372 G; 442 T; 4 other;

Alignment Scores:

Pred. No.: 1.04e-150 Length: 1618
Score: 1326.00 Matches: 257
Percent Similarity: 87.16% Conservative: 1
Best Local Similarity: 86.82% Mismatches: 0
Query Match: 97.57% Indels: 38
DB: 20 Gaps: 1

US-09-954-846-2 (1-258) x AA52251 (1-1618)

Qy 1 MetAlaValLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrp 20
Db 1 ATGGGGGCTTGGCACCTCTAATTGCTCTCGTGTATTCGGTGGCCGACTTTCAGATGG 60
Qy 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArg 40
Db 61 CTCGCCCAACCTTACTACCTTCTGCGGCCCTGCTCTGCTGCTCTCTACTCTGAGG 120
Qy 41 LysLeuProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysasp 60
Db 121 AAACCTGGCGCGCTCTGCCACGGTCTGCCACCCCAACGAGACGGTAACCCGTGTAC 180
Qy 61 PheAspTrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysasn 80
Db 181 TTTGACTGGAGAGAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAC 240
Qy 81 ArgArgSer----- 83
Db 241 CGCAGATCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGCC 300
Qy 83 ----- 83
Db 301 AACACAATCTTTCTCCGCTTGGATATTGCCATGGCGCTACTTTACATCACACTCTGC 360
Qy 84 ---MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102
Db 361 ATAGTGTTCCTGATGACGTGCAAAACCCCTATATATGGGCCCTGAGTATTAAGTAC 420
Qy 103 PheAsnAspLysThrIleAspGluGluLeuGluArgAspLysArgValThrTrpIleVal 122
Db 421 TTCATGATAAAACCATGTAGGAGAACTAGAACGGGACAGAGGGTCACTTGGATTGTG 480
Qy 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
Db 481 GAGTCTTTGGCAATTTGCTCTAATGACTGCCAATCATTTGGCCCTATCTATGCTCACCTC 540
Qy 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr 162
Db 541 TCCCTTAAATACAACTGTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACT 600
Qy 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
Db 601 GATGTTAGTAGCGGTACAAAGTCAGACATCACCCCTCACCAAGCACTCCCTACCTG 660
Qy 183 IleLeuPheGlnGlyLysGluAlaMetArgProGlnIleAspLysLysGlyArg 202
Db 661 ATCTGTTCCAAGGTGGCAAGGAGCAATGCGCGGCCACAGATTGACAAGAAAGACGG 720
Qy 203 AlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu 222
Db 721 GCTGCTCATGAGCTTCTCTGAGGAGAAATGTGATCCGAGAATTAACCTTAATGAGTA 780
Qy 223 TyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVal 242
Db 781 TACCAGGGGGCCCAAGAAACTATCAAGGCTGGAGACAATATCCCTGAGGACGAGCTGTG 840
Qy 243 AlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
Db 841 GCTTCAACCCACCACAGTGTCTAGATGGGGGAAAACAGAGAGATAAA 888

RESULT 5

QY 103 PheAsnAspLysThrIleAspGluLeuGluArgAspLysArgValThrTrpIleVal 122
Db 1210 TTCAATGATAAACCATTTGATGAGAACTAGAACGGGACAAAGAGGGTCACCTGGATTGTG 1151
QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
Db 1150 GAGTTCTTTGCCAATGGTCTAATGACTGCCAATCAATTTGCCCTATCTATGCTGACCTC 1091
QY 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPhelyLysValAspValGlyArgTyrThr 162
Db 1090 TCCTTAATACAACTACTACAGGGCTAAATTTTGGGAAGGTGGATGTGGACGCTATACT 1031
QY 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
Db 1030 GATGTTAGTACGGGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 971
QY 183 IleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArg 202
Db 970 ATCCTGTTCCAAAGGTGCAAGGAGGCAATGCGCGGCCACAGATTGACAAGAAAGGAGCG 911
QY 203 AlaValSerThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu 222
Db 910 GCTGTCATCGGACCTTCTCTGAGGAGAATGTGATCGGAGAATTAATTAATGAGCTA 851
QY 223 TyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVal 242
Db 850 TACCAGCGGCCCAAGAACTATCAAGGCTGGAGACAATATCCCTGGAGGACGCTGTG 791
QY 243 AlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
Db 790 GCITCAACCCCCACACAGTGTGATGGGGAACAAAGAGGATATA 743
RESULT 7
AAV34294
ID AAV34294 standard; DNA; 1699 BP.
XX
AC AAV34294;
XX
DT 29-JAN-1999 (first entry)
XX
DE Human secreted protein gene 9 clone HJBCD89.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasia.
XX
Homo sapiens.
PN WO9840483-A2.
XX
PD 17-SEP-1998.
XX
PF 12-MAR-1998; 98WO-US04858.
XX
PR 19-DEC-1997; 97US-0068368.
PR 14-MAR-1997; 97US-0040710.
PR 14-MAR-1997; 97US-0040762.
PR 30-MAY-1997; 97US-0048100.
PR 30-MAY-1997; 97US-0048189.
PR 30-MAY-1997; 97US-0048357.
PR 30-MAY-1997; 97US-0050934.
PR 08-JUN-1997; 97US-0048970.
PR 05-SEP-1997; 97US-0057765.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ferrie AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;
PI Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;

PI Wei YF, Young PE, Zeng Z;
XX WPI; 1998-520811/44.
DR P-PSDB; AAW75204.
XX
PT Isolated human poly-nucleotide(s) encoding secretory peptide(s) -
PT used to develop products for the diagnosis and treatment of e.g.
PT inflammation, cancers, CNS disorders or immune system disorders
XX
XX Claim 1; Page 121-122; 20pp; English.
XX
CC This sequence represents a nucleic acid molecule which encodes a
CC secreted human protein. The gene number, and the clone it is derived
CC from, are detailed in the descriptor line. The gene can be used to
CC generate fusion proteins by linking to the gene to a human immunoglobulin
CC Fc portion (e.g. AAV34277) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 28 novel genes and their fragments (nucleic
CC acid sequences: AAV34286-V34325; amino acid sequences AAV75196-W75235)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 28
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAV34286 for described uses).
XX
SQ Sequence 1699 BP; 465 A; 400 C; 383 G; 449 T; 2 other;

Alignment Scores:
Pred. No.: 3,41e-150 Length: 1699
Score: 1322.00 Matches: 256
Percent Similarity: 86.82% Conservative: 1
Best Local Similarity: 86.49% Mismatches: 1
Query Match: 97.28% Indels: 38
DB: 19 Gaps: 1
US-09-954-846-2 (1-258) x AAV34294 (1-1699)
QY 1 MetAlaValLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrp 20
Db 30 ATGCGGGTCTTGGCACCTCTAATGCTCTCGTGATTCTGGTGGCGGCGATTTCAGATGG 89
QY 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArg 40
Db 90 CTCGCCCAACCTTACTACTCTCTGCGCCCTGCTCTGCTGCTCTCTACTCTGCTGAGG 149
QY 41 LysLeuProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
Db 150 AAATCGCGCGCTCTGCCACGGTCTGCCACCGGCTGCCACCGGAGGAGGTAACCCGTGTGAC 209
QY 61 PheAspTrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsn 80
Db 210 TTTGACTGGAGAGAAGTGGAGATCTCTGATGTTTCTCAGTGCCATGTGATGATGAAGAC 269
QY 81 ArgArgSer-----
Db 270 CGCAGATCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGCC 329
QY 83 -----
Db 330 AACACAATCTTTCTCCGCTTGGATATTCGCATGGGCCCTACTTTACATCACACTCTGC 389
QY 84 ---MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102
Db 390 ATAGTGTCTGATGACGTGCAACACCCCTCTATATATGCGCCCTGAGTATATCAAGTAC 449
QY 103 PheAsnAspLysThrIleAspGluGluLeuGluArgAspLysArgValThrTrpIleVal 122
Db 450 TTCAATGATAAAACCATTTGATGAGGAAGTACAGAACGGGACAGAGGGTCACTTGGATTGTG 509
QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
Db 1210 TTCAATGATAAACCATTTGATGAGAACTAGAACGGGACAAAGAGGGTCACCTGGATTGTG 1151

510 GAGTCTCTTGGCAATGGTCTAATGACTGCAATCATTTGGCCCTATCTATGCTGACCTC 569
143 SerLeuLysThrAsnCysThrGlyLeuAsnPhcGlyLysValAspValGlyArgTyrThr 162
570 TCCCTTAAATACAACTGTACAGGCTAAATTTTGGGAAGTGGATGTTGGACGCTATACT 629
163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
630 GATGTTAGTACGGGTACAAAGTGAGCACATACCCCTCACCAGCAACTCCCTACCTG 689
183 IleLeuPheGlnGlyLysGluAlaMetArgProGlnIleAspLysLysGlyArg 202
690 ATCCTCTTCCAGTGGCAAGGAGCAATCGCGGCGCACAGATTGACAAAGAGGACGG 749
203 AlaValSerThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu 222
750 GCTGCTCATGGACCTCTCTGAGGAGATGTGATCCGAGATTTAACTTAAATGAGCTA 809
223 TyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVal 242
810 TACCAGCGGCGCAAGAACTATCAAGGCTGGAGACAAATATCCTTGAGGAGCAGCTGTG 869
243 AlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
870 GNTTCAACCCCAACACAGTGTGATGGGGAACAAAGAGGATAA 917

RESULT 8

AAF72409

ID AAF72409 standard; cdna; 1621 BP.

XX AC AAF72409;

XX DT 24-APR-2001 (first entry)

XX DE Human PRO270 cdna.

Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
antiParkinsonian neurotropic; neuroprotective; vulnerary; cardiant;
antiangiogenic; vasotropic; antiasthmatic; anticirneumatic; cancer;
antiarthritic; antinfertility; antidiabetic; antiviral; diabetes;
ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
ischaemia; inflammation; ss.

OS Homo sapiens.

WO200104311-A1.

18-JAN-2001.

22-FEB-2000; 2000WO-US04414.

07-JUL-1999; 99US-0143048.

26-JUL-1999; 99US-0145698.

28-JUL-1999; 99US-0146222.

08-SEP-1999; 99WO-US20594.

13-SEP-1999; 99WO-US20944.

15-SEP-1999; 99WO-US21090.

15-SEP-1999; 99WO-US21547.

29-OCT-1999; 99WO-US23089.

29-NOV-1999; 99WO-US28214.

30-NOV-1999; 99WO-US28313.

16-DEC-1999; 99WO-US30095.

20-DEC-1999; 99WO-US30911.

20-DEC-1999; 99WO-US30999.

05-JAN-2000; 99WO-US00219.

(GETH) GENENTECH INC.

PA Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI.

XX WPI: 2001-081051/09.
DR P-PSDB: AAB80248.

XX Sixty one nucleic acids encoding PRO polypeptides which are useful in
PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
PT squamous cell carcinoma) and neurodegenerative diseases (e.g.
PT Alzheimer's disease)

XX Claim 2; Fig 75; 393pp; English.

XX The present sequence is one of sixty one nucleic acids encoding novel
CC secreted and transmembrane PRO polypeptides. The PRO polypeptides are
CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
CC squamous cell carcinoma), gastrointestinal disorders (e.g.
CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.
CC endometrial bleeding angiogenesis, ischaemias such as coronary
CC ischaemia, atherosclerosis); inflammatory disorders (e.g. asthma,
CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
CC diabetes and retinal disorders such as retinitis pigmentosum.
CC The PRO nucleic acids have applications in molecular biology, including
CC use as hybridization probes, and in chromosome and gene mapping.

XX Sequence 1621 BP; 407 A; 394 C; 374 G; 442 T; 4 other;
SQ

Alignment Scores:

Pred. No.: 3.94e-149 Length: 1621
Score: 1313.00 Matches: 257
Percent Similarity: 86.87% Conservative: 1
Best Local Similarity: 86.53% Mismatches: 0
Query Match: 96.62% Indels: 39
DB: 22 Gaps: 1

US-09-954-846-2 (1-258) x AAF72409 (1-1621)

QY 1 MetAlaValLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrp 20

Db 3 ATGGCGGTCTGGCACCTCTAAATGCTCTCGTGTATTGGTGGCGGCGGATTTTCAGATGG 62

QY 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuValArg 40

Db 63 CTGCGCCCAACCTTACTACTCTCTGCGGCCCTCTCTGCTGCTCTCTCTACTCTGAGG 122

QY 41 LysLeuProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60

Db 123 AAATCGCGCGGCTCTGCCACGGTCTGCCACCCACCGAAGACGGTAACCCGTGTGAC 182

QY 61 PheAspTrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsn 80

Db 183 TTTGACTGGAGAGAAAGTGGAGATCCTCATGTTTCTCAGTGCCTTGTGATGATGAAGAAC 242

QY 81 ArgArgSer 83

Db 243 CGCAGATCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGCC 302

QY 83 83

Db 303 AACACAATCTTTTCTCGCTGGATATTCGCATGGCCCTTACTTTTACATCACACTCTGC 362

QY 84 ---MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102

Db 363 ATAGTGTCTGTGATGACGTCAACACCCCTTATATGGGCCCTAGTATATCAAGTAC 422

QY 103 PheAsnAspLysThrIleAspGluLeuGluArgAspLysArgValThrTrpIleVal 122

Db 423 TTCAATGATAAACCATTGATGAGAACTAGACGGGACAGAGGGTCACCTTGGATTGTG 482

QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142

Db 483 GAGTCTCTTCCCAATGGTCTAATGACTGCCAATCATTTGCCCTATCTATGCTGACCTC 542

QY 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr 162

```

Db 543 TCCCTTAATACTGACAGGGCTAAATTTGGGAAGGTGGATGTTGGACGCCTACT 602
Qy 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
Db 603 GATGTTAGTACGCGTACAAAGTGAGACATCACCCCTCACCAAGCAACTCCCTACCCCTG 662
Qy 183 IleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArg 202
Db 663 ATCTCTGTTCCAGGTGGCAAGGAGCAATGCGGCGCCACAGATTGACAAGAAAGGACGG 722
Qy 203 AlaValSerTTP-ThrPheSerGluGluAsnValIleArgGluPheAsnLeuAspGluLe 222
Db 723 CTTGCTCATGGACCTTCTCTGAGCAGAAAGTGATCCGAGAAATTAATTAATTAAGCT 782
Qy 222 uTyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVa 242
Db 783 ATACAGCGGGCCAAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGT 842
Qy 242 lAlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
Db 843 GGCTTCAACCCCAACACAGTGTCTAGATGGGAAACAAAGAGGATAAA 891
JUL 9
34316
ID AAV34316 standard; DNA; 1652 BP.
XX
AC AAV34316;
XX
DT 29-JAN-1999 (first entry)
DE Human secreted protein gene 9 clone HJBCD89.
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO9840483-A2.
XX
PD 17-SEP-1998.
XX
PE 12-MAR-1998; 98WO-US04858.
XX
Pr 19-DEC-1997; 97US-0068368.
Pr 14-MAR-1997; 97US-0040710.
Pr 14-MAR-1997; 97US-0040762.
Pr 30-MAY-1997; 97US-0048100.
Pr 30-MAY-1997; 97US-0048189.
Pr 30-MAY-1997; 97US-0048357.
Pr 30-MAY-1997; 97US-0050934.
Pr 06-JUN-1997; 97US-0048970.
Pr 05-SEP-1997; 97US-0057765.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ferrie AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;
PI Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;
PI Wei YF, Young PE, Zeng Z;
XX
DR WPI; 1998-520811/44.
DR P-PSDB; AAW5226.
XX
PT Isolated human poly:nucleotide(s) encoding secretory peptide(s) -
PT used to develop products for the diagnosis and treatment of e.g.
PT inflammation, cancers, CNS disorders or immune system disorders
XX

```

```

PS Claim 1; Page 143-144; 201pp; English.
XX
CC This sequence represents a nucleic acid molecule which encodes a
CC secreted human protein. The gene number, and the clone it is derived
CC from, are detailed in the descriptor line. The gene can be used to
CC generate fusion proteins by linking to the gene to a human immunoglobulin
CC Fc portion (e.g. AAV34277) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 28 novel genes and their fragments (nucleic
CC acid sequences: AAV34286-V34325; amino acid sequences AAW5196-W5235)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 28
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAV34286 for described uses).
XX
SQ Sequence 1652 BP; 423 A; 396 C; 379 G; 448 T; 6 other;

Alignment Scores:
Pred. No.: 4,06e-149 Length: 1652
Score: 1313.00 Matches: 254
Percent Similarity: 87.03% Conservative: 1
Best Local Similarity: 86.69% Mismatches: 0
Query Match: 96.62% Indels: 38
DB: 19 Gaps: 1

US-09-954-846-2 (1-258) x AAV34316 (1-1652)

Qy 4 LeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrpLeuAlaGln 23
Db 1 TTGGCACCTCTAATTGCTCTCTGCTATTCGGTGGCGGACTTTCAGATGGGTGCGCCCAA 60
Qy 24 ProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArgLysLeuPro 43
Db 61 CCTTACTACTCTCTGCGCCCTGCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Qy 44 ProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAspPheAspTrp 63
Db 121 CCGCTCTGCGACGGTCTGCCACCCACCGAAGCGGTAAACCGCTGCTGCTGCTGCTGCTG 180
Qy 64 ArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsnArgArgSer 83
Db 181 AGAGAAGTGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCCGACATCC 240
Qy 83 ----- 83
Db 241 ATCACTGTGGACCAACATATAGGCAACATTTTCATGTTTAGTAAGTGGCCCAACAATT 300
Qy 84 -----MetPhe 85
Db 301 CTTTCTTCCGCTTGGATATTCGCATTCGCGCTACTTTACATCACACTCTGCATAGTGTTC 360
Qy 86 LeuMetThrCysLysProLeuTyrMetGlyProGluTyrIleLysTyrPheAsnAsp 105
Db 361 CTGATGACGTGCAAAACCCCTATATATGCGSCCTGAGTATATCAAGTACTCAATGAT 420
Qy 106 LysThrIleAspGluLeuGluArgAspLysArgValThrTrpIleValGluPhePhe 125
Db 421 AAACACATTGATCAGGAAGTAAAGCGGACAGAGGGTCACTTGGATTTGGAGTTCTTT 480
Qy 126 AlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeuSerLeuLys 145
Db 481 GCCAATTGCTCTAATGACTGCCCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAA 540
Qy 146 TyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThrAspValSer 165
Db 541 TACAACCTGTACAGGGCTAAATTTTGGGAAGGTGGATTTGGACGCTACTACTGTTAGT 600
Qy 166 ThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeuIleLeuPhe 185
Db 601 ACGCGGTACAAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTTACCTGTATCTGTTTC 660

```

QY 186 GlnGlyGlyLysGluAlaMetArgProGlnIleAspLysLysGlyArgAlaValSer 205
 Db 661 CAAGGTGGCAGGAGCAATCGCGCGCCACAGATTGACAAAGACGGGTGTCTCA 720
 QY 206 TrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeuTyrGlnArg 225
 Db 721 TGGACCTTCTCTCAGGAGAGATGTATCCGAGAAATTTAACTTAATGAGCTATACCCAGCG 780
 QY 226 AlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProValAlaSerThr 245
 Db 781 GCCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACC 840
 QY 246 ProThrThrValSerAspGlyGluAsnLysLysAspLys 258
 Db 841 CCCACCACAGTGTCTAGATGGGAAACAAAGAGGATATA 879

ULT 10

...L49614
 ID ABL49614 standard; DNA: 1659 BP.

XX ABL49614;

AC 01-JUN-2002 (first entry)

DT Prostaglandin E2 (PGE2) encoding DNA.

DE Prostaglandin E2; PGE2; stomach cancer; gene; ds.

XX Unidentified.

XX Key Location/Qualifiers

FT CDS 31..921

FT /*tag= a

FT /product= "prostaglandin E2"

XX KR2001081233-A.

XX 29-AUG-2001.

XX 11-FEB-2000; 2000KR-0006416.

XX 11-FEB-2000; 2000KR-0006416.

XX (NAHE-) NAT INST HEALTH IN KOREA.

XX Kim GC, Nam MJ, Park MS;

XX WPI; 2002-136917/18.

XX P-PSDB; ABB06300.

XX Novel gene expressed by prostaglandin E2 (PGE2) in stomach cancer cell

PT line snu-1 and its amino acid sequence -

XX Disclosure; Page 12; 12pp; Korean.

XX The present sequence encodes prostaglandin E2 (PGE2) which is expressed

CC in the stomach cancer cell line SNU-1. The PGE2 gene can be used in the

XX treatment of stomach cancer.

XX Sequence 1659 BP; 426 A; 404 C; 380 G; 448 T; 1 other;

XX Alignment Scores:

Pred. No.: 1.65e-148 Length: 1659

Score: 1308.00 Matches: 255

Percent Similarity: 86.49% Conservative: 1

Best Local Similarity: 86.15% Mismatches: 2

Query Match: 96.25% Indels: 38

DB: 24 Gaps: 1

Db 31 ATGCGGCTTGGCACCTCTAATTGCTCTCGTATTTCGGTCCGCGACTTTCAGATGG 90
 QY 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArg 40
 Db 91 CTCGCCCAACCTTACTACCTTCTGTGGCCCTGTCTCTGTCTGTCTCTACTCTGTAGG 150
 QY 41 LysLeuProLeuLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
 Db 151 AAACCTCGCGCGCTCTGCCACGGTCTGCCACCCCAACGCGAAGAGGTAACCCGTGTAC 210
 QY 61 PheAspTrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsn 80
 Db 211 TTTGACTGGAGAGAAGTGGAGATCTGTATGTTTCTCAGTGCATTTGTGATGATGAAGAC 270
 QY 81 ArgArgSer----- 83
 Db 271 CGCAGATCCATCCTACTGTGGAGCAACATATAGGCAACATTTTCATGTTTGTAGTAAAGTGCC 330
 QY 83 ----- 83
 Db 331 AACACAATCTTTTCTCCGCTTGGATATTCGCATGGGCCCTACTTTACATCACACTCTGC 390
 QY 84 ---MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102
 Db 391 ATAGTGTCTGTGATGAGTGCACAAACCCCTATATATATGCGCCTGAGTATATCAAGTAC 450
 QY 103 PheAsnAspLysThrIleAspGluLeuGluArgAspLysArgValThrTrpIleVal 122
 Db 451 TTCAATGATAAACCATTTGATGAGGAACTAGAACGGGACAAAGAGGGTCTACTTGGATTGTG 510
 QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
 Db 511 GAGTTCTTGGCAATTTGGTCTAAATGACTGCCAATCATTTGCCCTTATCTATGCTGACCTC 570
 QY 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr 162
 Db 571 TCCCTTAAATACAACTGTACAGGCTAAATTTGGGAAGGTGGATGTTGGACGCTATACT 530
 QY 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
 Db 631 GATGTTAGTACGCGTACAAAGCGAGCACATCACCCCTCACCAAGCAACTCCCTACCCCTG 690
 QY 183 IleLeuPheGlnGlyLysGluAlaMetArgProGlnIleAspLysLysGlyArg 202
 Db 691 ATCCTGTCCAAAGTGGCAAGGAGCAATGCGCGGCCACAGATTGACAAGAAAGGACGG 750
 QY 203 AlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu 222
 Db 751 GCTGCTCACGGACCTTCTCTGAGAGAGATGTGATCCGAGAAATTAACTTAATGAGCTA 810
 QY 223 TyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVal 242
 Db 811 TACCAGCGGGCCAAAGAACTATCAAGGCTGGAGACAATATCCCTCTGAGGAGCAGCCTGTG 870
 QY 243 AlaSerThrProThrValSerAspGlyGluAsnLysLysAspLys 258
 Db 871 GCTTCAACCCCCACACAGTGTGATGGGGGAAACAAAGAGGATAAA 918

RESULT 11

AAF93762

ID AAF93762 standard; cDNA; 1631 BP.

XX AAF93762;

AC AAF93762;

XX 23-MAY-2001 (first entry)

DT Human cDNA encoding a membrane or secretory protein clone PSEC0045.

XX Human; secretory protein; membrane protein; vaccine; gene therapy;

XX rheumatoid arthritis; diabetes; ss.

XX Homo sapiens.

OS

QY 1 MetAlaValLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrp 20

XX EP1067182-A2.
 XX 10-JAN-2001.
 XX 07-JUL-2000; 2000EP-0114090.
 XX 08-JUL-1999; 93JP-0194179.
 PR 11-JAN-2000; 2000JP-0118775.
 PR 02-MAY-2000; 2000JP-0183766.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K,
 PI WPI; 2001-093989/11.
 XX P-PSDB; AAB88335.
 DR Nucleic acids encoding secretory proteins/membrane proteins, useful in
 PT gene therapy or as candidate target molecules in drug development -
 XX Claim 1; SEQ ID 37; 609pp + CD ROM; English.
 vv This invention relates to nucleic acid sequences AAF93744 - AAF93916
 CC which encode human secretory or membrane proteins represented by
 CC AAB88317 - AAB88419. Included in the invention are primers
 CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
 CC cDNA sequences of the invention. The invention also includes methods for
 CC the production of antibodies directed against the proteins, and cDNA
 CC sequences, which can be used in vaccines. The polynucleotide sequences
 CC can be used in gene therapy. The polynucleotide sequences and the
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate secretory
 CC protein/membrane protein expression. The nucleic acids and complementary
 CC sequences may also be used as DNA probes in diagnostic assays
 CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples. They may also be
 CC used to study the expression and function of secretory proteins/membrane
 CC polypeptides and their role in metabolism. The polypeptides may be used
 CC as antigens in the production of antibodies against them and in assays to
 CC identify modulators (agonists and antagonists) of expression and
 CC activity. The antibodies and antagonists may also be used as therapeutic
 CC agents to down regulate expression and activity. The antibodies may also
 CC be used as diagnostic agents for detecting the presence of the
 CC polypeptides in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA)). Examples of diseases which may be treated include rheumatoid
 CC arthritis and diabetes.
 XX
 SQ Sequence 1631 BP; 411 A; 397 C; 378 G; 445 T; 0 other;

Alignment Scores:
 d. No.: 2e-147 Length: 1631
 re: 1299.00 Matches: 256
 Percent Similarity: 86.82% Conservative: 1
 Best Local Similarity: 86.49% Mismatches: 1
 Query Match: 95.58% Indels: 39
 DB: 22 Gaps: 1

US-09-954-846-2 (1-258) x AAF93762 (1-1631)

QY 1 MetAlaValLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTTP 20
 Db 16 ATGGCGGTCTGGACCTTAATGCTCGGTGATTCGGTCCGGGACTTTCACCATGG 75
 QY 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuValArg 40
 Db 76 CTCGCCCAACCTTACTACCTCTGTGGCCCTGCTCTGCTGCTTCTACTCGTGAGG 135
 QY 41 LysLeuProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
 Db 136 AAATCGCCCGCTCTGCGACGCTGCTGCCACCCACCGAGACGTAACCCGCTGTGAC 195
 QY 61 PheAspTyrArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsn 80

Db 196 TTTCAGTGGAGAGAGAGTGGAGATCCTGATGTTTCTCAGTGGCCATTGTGATGATGAAGAAC 255
 QY 81 ArgArgSer-----
 Db 256 CGCAGATCCATCCTCTGTGGAGCAACATATAGGCAACATTTTTCATGTTTAGTAAAGTGCC 315
 QY 83 -----
 Db 316 AACACAATTCTTTCTCCGCTTGGATATTCGCATGGGCTACTTTTACATCACACTCTGC 375
 QY 84 ---MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102
 Db 376 ATAGTCTTCTGTATGAGTGCACAAACCCCTATATATATGGCCCTGAGTATATCAAGTAC 435
 QY 103 PheAsnAspLysThrIleAspGluGluLeuGluArgAspLysArgValThrTrrIleVal 122
 Db 436 TTCAATGATAAAACCATTTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTG 495
 QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
 Db 496 GAGTTCTTTGCCAATTTGGTCTAATGACTGCAATCATTTTGCCTCTATCTATGCTGACCTC 555
 QY 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr 162
 Db 556 TCCCTTAATATACAACTGTACAGGGCTAAATTTGGGAGGTGGATGTGGACCTATACT 615
 QY 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
 Db 616 GATGTTAGTAGCGGTACAAAGTGACACATCACCTCCACCAAGCAACTCCCTACCTG 675
 QY 183 IleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArg 202
 Db 676 ATCTCTTCCAAGTGGCAAGGAGCAATGCGGGCCACAGATTGACAAAGAAGACGG 735
 QY 203 AlaValSerTyrThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu 222
 Db 736 GCTGTCTCATGGACCTTCTCTGAGGAGATGTCATCCGAGAATTAACTTAATGAGCTA 795
 QY 223 TyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVal 242
 Db 796 TACCAGCGGCCCAAGAAACTATCAAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTG 855
 QY 243 AlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
 Db 856 GC-TCAACCCCCCACCACAGTGTGAGTGGGGGAAACAAGAAGGATAAA 902

RESULT 12
 AA242527
 ID AA242527 standard; cDNA; 774 BP.
 XX
 AC AA242527;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human 5' EST isolated from a cDNA library SEQ ID NO:286.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
 KW gene therapy; chromosome mapping; upstream regulatory sequence;
 KW forensic; location; development; protein synthesis; stability;
 KW regulation; identification; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9953051-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 09-APR-1999; 99WO-1B00712.
 XX
 PR 09-APR-1998; 98US-0057719.
 PR 28-APR-1998; 98US-0069047.
 XX

PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-038446/03.
 XX P-PSDB; AAY64913.
 DR
 XX
 PT Novel secreted protein 5' expressed sequence tag sequences used in
 PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
 XX
 XX Claim 1; Page 324; 837pp; English.
 XX
 CC AAZ42265 to AAZ43075 represent novel 5' expressed sequence tag (EST)
 CC sequences, corresponding to human secreted proteins. AAY64651 to
 CC AAY65438 represent the EST-related proteins corresponding to AAZ42265 to
 CC AAZ43052. The 5' ESTs can be used for producing secreted human gene
 CC products. They can be used to identify and isolate 5' untranslated
 CC regions (UTRs) and upstream regulatory regions which control the
 CC location, development stage, rate, and quantity of protein synthesis, as
 CC well as stability of mRNA. The ESTs are also useful as probes for
 CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
 CC also be used in forensic procedures to identify individuals, or in
 CC diagnostic procedures to identify individuals having genetic diseases
 CC resulting from abnormal gene expression. The products may also be used in
 CC gene therapy protocols. The nucleic acids encoding signal peptides can be
 CC used for directing extracellular secretion of a polypeptide or the
 CC insertion of a polypeptide into a membrane, or importing a polypeptide
 CC into a cell. The proteins encoded by the EST sequences may be useful in
 CC treating a variety of human conditions. Secreted proteins have
 CC therapeutic value, and the identification of new secreted proteins is
 CC valuable. AAZ42249 to AAZ42264 and AAY64644 to AAY64650 represent
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 774 BP; 181 A; 203 C; 193 G; 195 T; 2 other;

Alignment Scores:

Pred. No.: 7.5e-111 Length: 774
 Score: 994.00 Matches: 193
 Percent Similarity: 82.91% Conservative: 1
 Best Local Similarity: 82.48% Mismatches: 2
 Query Match: 73.14% Indels: 38
 DB: 21 Gaps: 1

US-09-954-846-2 (1-258) x AAZ42527 (1-774)

1 MetAlaValLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrrp 20
 |||||
 71 ATGGCGGTCTGGCACTCTAATGTCTCTGCTGTTATTCGGTGGCGGACITTCAGATGG 130
 |||||
 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuValArg 40
 |||||
 131 CTGCGCCCACTTACTACCTCTGTCTGCGGCTGTCTCTGCTGCTCTCTACTCTGAGG 190
 |||||
 41 LysLeuProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
 |||||
 191 AATCTCGCGCGGCTGTCCACGGTCTGCCACCCACCGAAGCGGTAACCCGTGTGAC 250
 |||||
 61 PheAspTrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsn 80
 |||||
 251 TTTGACTGGAGAGAGTGGAGATCCTGTGTTCTCAGTGCCATTTGTGATGATCAAGAAC 310
 |||||
 81 ArgArgSer----- 83
 |||||
 311 CGCAGATCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGCC 370
 |||||
 83 ----- 83
 371 AACACAAATCTTTTCTCCGCTGGATPATTGCGATGGCGCTACTTTTACATCACTGCG 430
 |||||
 84 ---MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102
 |||||
 431 ATAGTGTCTGATGAGTGCAACACCCCTATATATATGGCCCTGATATATCANGTAC 490

QY 103 PheAsnAspLysThrIleAspGluLeuGluArgAspLysArgValThrTripleVal 122
 |||||
 Db 491 TTCATGATAAACACCATGTAGAGAACTAGACGGGACAGAGGGTCACTTGGATTGTG 550
 |||||
 QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
 |||||
 Db 551 GAGTTCITTTGGCAANTGGTCTAATGACTGCCAATCATTTGGCCCTATCTATGCTGACCTC 610
 |||||
 QY 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr 162
 |||||
 Db 611 TCCCTTAAATACAACTGTACAGGGCTAAATTTTGGAAAGGTGGATGTTGGAGCGCTATACT 670
 |||||
 QY 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
 |||||
 Db 671 GATGTTAGTACGGGTACAAAGTGAGCACATCACCCCTCACCAAGCAAACTCCCTACCTG 730
 |||||
 QY 183 IleLeuPheGlnGlyLysGluAlaMetArgArgProGln 196
 |||||
 Db 731 ATCCTGTTCCAAGTGGCAAGGAGGCAATGCGGGCGGCACAG 772
 |||||
 RESULT 13
 AAF93955
 ID AAF93955 standard; DNA; 726 BP.
 XX
 AC AAF93955;
 XX
 DT 23-MAY-2001 (first entry)
 XX
 DE Primer specific for DNA encoding secretory/membrane protein SEQ ID 389.
 XX
 KW Human; secretory protein; membrane protein; vaccine; gene therapy;
 KW rheumatoid arthritis; diabetes; PCR primer; ss.
 XX
 OS Synthetic.
 XX
 PN EP1067182-A2.
 XX
 PD 10-JAN-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114090.
 XX
 PR 08-JUL-1999; 95JP-0194179.
 PR 11-JAN-2000; 2000JP-0118775.
 PR 02-MAY-2000; 2000JP-0183766.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
 XX WPI; 2001-093989/11.
 XX
 PT Nucleic acids encoding secretory proteins/membrane proteins, useful in
 PT gene therapy or as candidate target molecules in drug development -
 XX
 PS Claim 4; SEQ ID 389; 609pp + CD ROM; English.
 XX
 CC This invention relates to nucleic acid sequences AAF93744 - AAF93916
 CC which encode human secretory or membrane proteins represented by
 CC AAB88317 - AAB88419. Included in the invention are primers
 CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
 CC cDNA sequences of the invention. The invention also includes methods for
 CC the production of antibodies directed against the proteins, and cDNA
 CC sequences, which can be used in vaccines. The polynucleotide sequences
 CC can be used in gene therapy. The polynucleotide sequences and the
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate secretory
 CC protein/membrane protein expression. The nucleic acids and complementary
 CC sequences may also be used as DNA probes in diagnostic assays
 CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples. They may also be
 CC used to study the expression and function of secretory proteins/membrane
 CC polypeptides and their role in metabolism. The polypeptides may be used
 CC as antigens in the production of antibodies against them and in assays to

Db 161 TGGCTCGCCCAACCTTACTACCTTCTGTGCGCCCTGCTCTCTGCTGCCCTTCCCTACTCG 220
Qy 39 aArgLysLeuProLeuCysHisGlyLeuProThrGln-ArgGluAspGlyAsn-Pr 58
Db 221 TGAGGAACATGCGCGCTCTGCGACGCTCTGCCACCCACCAACCGAAGACGGTAACCC 280
Qy 58 OCysAspPheAspTrpArgGlu-ValGluIleLeuMetPheLeuSerAlaIleValMetM 78
Db 281 GTCTGACTTTTGACTGGAGAAAGTGAGATCCTGTGATTTCTCAGTGCCCATTTGTATGA 340
Qy 78 etLysAsn-ArgArgSer----- 83
Db 341 TGAGAACCCGACATCCATCCTGTGGCAACATATAGCAACATTTTCATGTTTAAG 400
Qy 83 ----- 83
401 TAAAGTGGCCCAACACAAATCTTTCTTCCGCTTGGATATTCGATGGCGCTACTTTAC 460
44 MetPheLeuMetThrCysLys-----ProProLeuTyrMetGly 96
Db 461 ATCACACTGCTAGTATGTTCTGATGACGTGCAACCCCTCCCTCCCA-TATATGGC 519
Qy 97 ProGluTyrIleLysTyrPheAsn-AspLysThrIleAspGluLeuGluArgAspLys 116
Db 520 CCTGATATATCAAGTACTTCAATGGATAAACCTTTGATGAGCACTAGAACGGGACAA 579
Qy 116 sArgValThrTrp-IleValGluPhePheAlaAsnTrpSerAsnAspCysGlnSerPhe 135
Db 580 CAGGTCAATTTGGAATTTGGAGTCTTTGGCAATTTGGTCTAATGACTGGCAATCAATT 639
Qy 136 AlaProIleTyrAlaAspLeuSerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLys 155
Db 640 GCCCTATCTATGCTGACCTCTCCCTTAAATACAACTGTACAGGCTAAATTTTGGGAG 699
Qy 156 ValAspValGlyArgTyrThrAspValSerThrArgTyrLysValSerThrSerProLeu 175
Db 700 GTGGATGTTGGACGTACTACTGATGTAGTACGGCGTACAAAGTGAGCACATCACCCCTC 759
Qy 176 ThrLys-GlnLeuProThrLeu-IleLeuPheGlnGlyGlyLysGluAlaMet---ArgA 194
Db 760 ACCAAGGCAACTCCCTACCTGGATCTCTTCCAAAGTGGGCAAGGAGGCAATCGGGC 819
Qy 194 rgProGlnIle---AspLysLysGlyArg-AlaVal-SerTrpThrPhe----- 208
Db 820 GGCCACAGATTTGCACAAAGGAAGGAGCGGGCTGTTTCATGGGACCTCTCTGAGGTAC 879
Qy 209 -----SerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeuTyrGlnAr 225
Db 880 CTGAAGGAAGGCGGAGGAGATGTATCCGAGAAATTAATTAATGAGCTATACCGCG 939
Qy 225 gAlaLysLysLeuSerLysAlaGlyAspAsnIle-ProGluGluGlnProValAlaSerT 245
Db 940 GGCAAGAACTATCAAGGCTGGAGACAATATCCCTGGAGGAGCAGGCTGTGGCTTCAA 999
Qy 245 hrProThrValSerAsp-GlyGluAsnLysLysAspLys 258
Db 1000 CCCCCACCAAGTGTGATGGGGGAAACAAAGGAAGATAAA 1041

RESULT 15

AAAA1964
ID AAA41964 standard; cDNA; 439 BP.
XX
AC AAA41964;
XX
XX 21-AUG-2000 (first entry)
XX
DE Human secreted expressed sequence tag SEQ ID NO:704.

XX Human; mouse; xenopus; rat; secreted expressed sequence tag; sBST;
KW expressed sequence tag; EST; probe; chemotactic; proliferative;
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KW antiviral; antidiabetic; antiasthmatic; vulnery; antiparkinsonian;

KW antilucer; osteopathic; neuroprotective; nootropic; antipsoriatic;
KW cerebroprotective; anticonvulsant; antiepressant; gene therapy;
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW central nervous system disorder; Alzheimer's disease; stroke;
KW parkinson's disease; Huntington's disease; coagulation disorder;
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW tumour; infection; depression; psoriasis; ss.

XX Homo sapiens.

OS WO200021990-A1.

PN 20-APR-2000.

XX 15-OCT-1999; 99WO-US24205.

PF 15-OCT-1998; 98US-0104435.

XX (GENY) GENETICS INST INC.

PA Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M;

XX WPI; 2000-317937/27.

DR Isolated polynucleotides, and encoded proteins, comprising secreted
XX expressed sequence tags (SESTs), useful for treating various disorders
XX such as autoimmune, infectious, and central nervous system disorders -

Claim 1; Page 315; 618pp; English.

XX AAA41261 to AAA43419 represent specifically claimed secreted expressed
CC sequence tags (SESTs), isolated from human, mouse, xenopus and rat
CC tissue sources. The SESTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:

CC chemotactic; proliferative; immunomodulatory; haematopoietic;
CC cytokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC cytosstatic; antibacterial; antifungal; antiviral; antidiabetic;
CC antiasthmatic; vulnery; antilucer; osteopathic; neuroprotective;
CC nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC anticonvulsant; and antiepressant. The SESTs can be used for gene
CC therapy and in vaccines. The SESTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA43420 to AAA43425 represent linker variants which are given
CC in the exemplification of the present invention.

XX Sequence 439 BP; 116 A; 111 C; 98 G; 114 T; 0 other;

Alignment Scores:

Pred. No.: 5,05e-66 Length: 439
Score: 622.00 Matches: 115
Percent Similarity: 99.15% Conservative: 1
Best Local Similarity: 98.29% Mismatches: 1
Query Match: 45.77% Indels: 0
DB: 21 Gaps:

US-09-954-846-2 (1-258) x AAA41964 (1-439)

Qy 84 MetPheLeuMetThrCysLysProLeuTyrMetGlyProGluTyrIleLysTyrPhe 103

Db 83 GTGTCTCTGATGACGTGCAACCCCTATATATGCGGCTGTATATCAAGTACTTC 142

```
QY 104 AsnAspLysThrIleAspGluGluLeuGluArgAspLysArgValThrTrpIleValGlu 123
Db 143 RATGATAAAACCATTTGATGAGGAAGTACAGGACGAGAGGTCACCTGGATTGTGGAG 202
QY 124 PhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeuSer 143
Db 203 TTCCTTGCCAATTGGTCTAATGACTGCCCAATCATTTGCCCTATCTATGCTGACCTCTCC 262
QY 144 LeuLysTyrAsnCysThrClyLeuAsnPheGlyLysValaspValGlyArgTyrThrAsp 163
Db 263 CTTAAATACAACTGTACAGGGCTAAATTTGGGAAGGTGGATGTTGGACGCTATACTGAT 322
QY 164 ValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeuIle 183
Db 323 GTTAGTAGCGGTACAAAGTGAGCACAATCCCTCACCAGCACTCCCTACCCCTGATC 382
QY 184 LeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLys 200
Db 383 CTGTTCCAAGGTGGCAAGGAGCAATGCGGCGGCCACACAGATTGACAAATAAA 433
```

Search completed: July 12, 2003, 11:13:16
--h time : 191 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 11:06:22 ; Search time 135 Seconds
(without alignments)
3017.635 Million cell updates/sec

Title: US-09-954-846-2

Perfect score: 1359

Sequence: 1 MAVLAPLIALYVSPRLSRW.....EQPVASTPTVSDGENKKDK 258

~oring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1105431 seqs, 789497651 residues

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framep2n.model -DEV=xlh
-O=/cn2.1/USPTO.spool/US0954846/runat_09072003_102146_12185/app_query.fasta_1.455
-DB=PublishedApplications_NA -QFWT=fastap -SUFFIX=rnpb -MINWATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-NAXLEN=200000000 -USER=US0954846.0CGN_1_1_125_0runat_09072003_102146_12185
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

tabase : PublishedApplications_NA:*

1: /cn2.6/ptodata/1/pubpna/US07_NEW_PUB.seq:
2: /cn2.6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cn2.6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cn2.6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cn2.6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
6: /cn2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:
7: /cn2.6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cn2.6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cn2.6/ptodata/1/pubpna/US09_NEW_PUB.seq:
10: /cn2.6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cn2.6/ptodata/1/pubpna/US10_NEW_PUB.seq:
12: /cn2.6/ptodata/1/pubpna/US10_PUBCOMB.seq:
13: /cn2.6/ptodata/1/pubpna/US60_NEW_PUB.seq:
14: /cn2.6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1359	100.0	10	US-09-954-846-4 Sequence 4, Appl
2	1359	100.0	10	US-09-925-300-751 Sequence 751, App
3	1326	97.6	1620	9 US-09-905-291A-206 Sequence 206, App
4	1326	97.6	1620	9 US-09-902-853-206 Sequence 206, App

5	1326	97.6	1620	9	US-09-907-824-206	Sequence 206, App
6	1326	97.6	1620	9	US-09-907-841-206	Sequence 206, App
7	1326	97.6	1620	9	US-09-904-011-206	Sequence 206, App
8	1326	97.6	1620	9	US-10-174-590-31	Sequence 31, Appl
9	1326	97.6	1620	9	US-10-176-758-31	Sequence 31, Appl
10	1326	97.6	1620	9	US-10-175-737-31	Sequence 31, Appl
11	1326	97.6	1620	9	US-09-906-742-206	Sequence 206, App
12	1326	97.6	1620	9	US-10-173-706-31	Sequence 31, Appl
13	1326	97.6	1620	9	US-10-175-738-31	Sequence 31, Appl
14	1326	97.6	1620	9	US-10-175-752-31	Sequence 31, Appl
15	1326	97.6	1620	9	US-10-176-482-31	Sequence 31, Appl
16	1326	97.6	1620	9	US-10-176-757-31	Sequence 31, Appl
17	1326	97.6	1620	9	US-10-176-913-31	Sequence 31, Appl
18	1326	97.6	1620	9	US-10-180-552-31	Sequence 31, Appl
19	1326	97.6	1620	9	US-10-180-557-31	Sequence 31, Appl
20	1326	97.6	1620	9	US-09-906-838-206	Sequence 206, App
21	1326	97.6	1620	9	US-09-907-613-206	Sequence 206, App
22	1326	97.6	1620	9	US-09-907-942-206	Sequence 206, App
23	1326	97.6	1620	9	US-10-173-700-31	Sequence 31, Appl
24	1326	97.6	1620	9	US-10-174-572-31	Sequence 31, Appl
25	1326	97.6	1620	9	US-10-174-579-31	Sequence 31, Appl
26	1326	97.6	1620	9	US-10-174-582-31	Sequence 31, Appl
27	1326	97.6	1620	9	US-10-174-588-31	Sequence 31, Appl
28	1326	97.6	1620	9	US-10-175-739-31	Sequence 31, Appl
29	1326	97.6	1620	9	US-10-175-740-31	Sequence 31, Appl
30	1326	97.6	1620	9	US-10-175-743-31	Sequence 31, Appl
31	1326	97.6	1620	9	US-10-176-488-31	Sequence 31, Appl
32	1326	97.6	1620	9	US-10-176-492-31	Sequence 31, Appl
33	1326	97.6	1620	9	US-10-176-747-31	Sequence 31, Appl
34	1326	97.6	1620	9	US-10-176-750-31	Sequence 31, Appl
35	1326	97.6	1620	9	US-10-176-985-31	Sequence 31, Appl
36	1326	97.6	1620	9	US-10-176-987-31	Sequence 31, Appl
37	1326	97.6	1620	9	US-10-176-991-31	Sequence 31, Appl
38	1326	97.6	1620	9	US-10-176-992-31	Sequence 31, Appl
39	1326	97.6	1620	9	US-10-176-993-31	Sequence 31, Appl
40	1326	97.6	1620	9	US-10-184-658-31	Sequence 31, Appl
41	1326	97.6	1620	9	US-10-173-695-31	Sequence 31, Appl
42	1326	97.6	1620	9	US-10-173-697-31	Sequence 31, Appl
43	1326	97.6	1620	9	US-10-173-705-31	Sequence 31, Appl
44	1326	97.6	1620	9	US-10-174-576-31	Sequence 31, Appl
45	1326	97.6	1620	9	US-10-174-585-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-954-846-4
: Sequence 4, Application US/09954846
: Patent No. US20020102654A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: Corley, Neil C.
: Guegler, Karl J.
: Patterson, Chandra
: Baughn, Mariah R.
: TITLE OF INVENTION: THIOREDOXIN PROTEINS
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows
: SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/954,846
: FILING DATE: 17-Sep-2001
: CLASSIFICATION: <Unknown>

141	AspLeuSerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArg	160
463	GACCTCTCCCTTAAATACAACTGACAGGGCTAAATTTGGGAAGTGATGTTGGACGC	522
161	TyrThrAspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuPro	180
523	TATACTGATGTATTACCGGTACAAAGTGAGCACATCACCCCTCACCAAGCAACTCCCT	582
181	ThrLeuIleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLys	200
583	ACCTGATCCCTGTTCCAAAGTGGCAAGGAGGCAATCGCGGGCCACAGATTGACAAAGAA	642
201	GlyArgAlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsn	220
643	GGACGGGCTGTCTCATGGACCTTCTCTCAGGAGAAATGTATCCGAGAAATTTAACTTAAT	702
221	GluLeuTyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGln	240
703	GAGCTATACCAAGCGGGCCAAAGAACTATCAAGGGCTGGAGACAATATCCCTGAGGAGCAG	762
241	ProValAlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys	258
763	CTGTGTGGCTTCAACCCCAACCACTGTCTCAAGATGGGAAACCAAGAAGATATAA	816

RESULTS

```

RESOLUTION 3
;
; Sequence 206, Application US/09905291A
; Patent No. US20020160374A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Tumas, Daniel
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
;
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;
; TITLE OF INVENTION: Acids Encoding the Same
;
; FILE REFERENCE: 10466-14
;
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
;
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
;
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
;
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
;
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
;
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
;
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
;
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15

```

```

> PRIOR APPLICATION NUMBER: PCT/US99/21547
> PRIOR FILING DATE: 1999-09-15
> PRIOR APPLICATION NUMBER: PCT/US99/23089
> PRIOR FILING DATE: 1999-10-05
> PRIOR APPLICATION NUMBER: PCT/US99/28214
> PRIOR FILING DATE: 1999-11-29
> PRIOR APPLICATION NUMBER: PCT/US99/28313
> PRIOR FILING DATE: 1999-11-30
> PRIOR APPLICATION NUMBER: PCT/US99/28564
> PRIOR FILING DATE: 1999-12-02
> PRIOR APPLICATION NUMBER: PCT/US99/28565
> PRIOR FILING DATE: 1999-12-02
> PRIOR APPLICATION NUMBER: PCT/US99/30095
> PRIOR FILING DATE: 1999-12-16
> PRIOR APPLICATION NUMBER: PCT/US99/30911
> PRIOR FILING DATE: 1999-12-20
> PRIOR APPLICATION NUMBER: PCT/US99/30999
> PRIOR FILING DATE: 1999-12-20
> PRIOR APPLICATION NUMBER: PCT/US00/00219
> PRIOR FILING DATE: 2000-01-05
> NUMBER OF SEQ ID NOS: 423
> SEQ ID NO 206
> LENGTH: 1620
> TYPE: DNA
> ORGANISM: Homo sapiens
> FEATURE:
> NAME/KEY: modified_base
> LOCATION: (973)
> OTHER INFORMATION: a, t, c or g
> NAME/KEY: modified_base
> LOCATION: (977)
> OTHER INFORMATION: a, t, c or g
> NAME/KEY: modified_base
> LOCATION: (996)
> OTHER INFORMATION: a, t, c or g
> NAME/KEY: modified_base
> LOCATION: (1003)
> OTHER INFORMATION: a, t, c or g
> US-09-905-291A-206

Alignment Scores:
Pred. No.: 1,84e-175
Score: 1326.00
Percent Similarity: 87.16%
Best Local Similarity: 86.82%
Query Match: 97.57%
DB: 9
Length: 1620
Matches: 237
Conservative: 1
Mismatches: 0
Indels: 38
Gaps: 1

```

US-09-954-846-2 (1-258) x US-09-905-291A-206 (1-1620)	
Qy	1 MetAlaValLeuAlaProLeuLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrrp 20
Db	3 ATGGCGGTCTTGGCACTCTAATTCCTCTCGTGTATTGGTCCCGGACTTTTCACGATGG 62
Qy	21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArg 40
Db	63 CTCGCCAACCTTACTACCTTCTGTCGCCCTGCTCTCTGTCGCCTTCTACTCTCGTAGG 122
Qy	41 LysLeuProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
Db	123 AAACCTGGCGCGCTCTGCCAGGTCTGCCACCCCAACGCGAAGAGCGTAACCCGTGTGAC 182
Qy	61 PheAspTrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsn 80
Db	183 TTTGACTGGAGAGAAGTGGAGATCCTGTAGTTTCTCAGTGCCATTGTGTATGATGAAGAAC 242
Qy	81 ArgArgSer----- 83
Db	243 CGCAGATCCATCACTGTGGAGCAACATATAGCAACATTTTTCATGTTTAGTAAAGTGCC 302
Qy	83 ----- 83
Db	303 AACACAAATCTTTTCTTCGCGTTGGATATTGCGATGGGCCCTACTTTACATCACACTTCG 362

Db 123 AAACGGCCGCGCTGTCGCCAGGTGTCGCCACCCCAACGCGAAGACGGTAACCCGTGTGAC 182
Oy 61 PheAspTTPAqGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsn 80
|||||
Db 183 TTTGACTGGAGAGAGTGGAGATCTCTGATGTTCTCAGTGCCATTGTGATGATGAGAC 242
Oy 81 ArgArgSer----- 83
|||||
Db 243 CGCAGATCCATCACTGTGGAGCAACATATAGCAACATTTTCATGTTTAGTAAAGTGGCC 302
Oy 83 ----- 83
Db 303 AACACAAFTCTTTTCTCCGCTTGGATATTCGCATGGGCTACTTTACATCACTCTGC 362
Oy 84 ---MetPheLeuMetThrCysLysProPheLeuTyrMetGlyProGluTyrIleLysTyr 102
|||||
363 ATAGTGTCTCTGATGACGTGCAACCCCTATATATGGGCCCTGAGTATATCAAGTAC 422
Db 423 TTCAATGATAAACCATTTGATGAGGAACCTAGAACGGACAGAGGCTCACTTGGATTGG 482
Oy 103 PheAsnAspLysThrIleAspGluGluLeuGluArgAspLysArgValThrTrpIleVal 122
|||||
Db 423 TTCAATGATAAACCATTTGATGAGGAACCTAGAACGGACAGAGGCTCACTTGGATTGG 482
Oy 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrIleAspLeu 142
|||||
Db 483 GAGTCTTTGGCAATTTGCTTAATGACTGCCAATCATTTGCCCTATCTATCTGACCTC 542
Oy 143 SerLeuLysTyrAsnCysThrClyLeuAsnPhGlyLysValAspValGlyArgTyrThr 162
|||||
Db 543 TCCCTTAAATAACAAGTACAGGGCTAAATTTGGGAAGGTGGATTTGGAGCGCTATACT 602
Oy 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
|||||
Db 603 GATGTACTAGCGGTACAAGTGAGACATCACCCCTCACCAAGCACTCCCTACCTCG 662
Oy 183 IleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArg 202
|||||
Db 663 ATCCTGTTCAGGTGGCAAGGAGGCAATGGCGGCCACACAGATTGACAAGAAAGACGG 722
Oy 203 AlaValSerThrPheSerGluAlaValIleArgGluPheAsnLeuAsnGluLeu 222
|||||
Db 723 GCTGTCTCATGACCTTCTCGAGGAGATGTGATCCGAGAATTTAACTTAAATGAGCTA 782
Oy 223 TyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVal 242
|||||
783 TACCAGCGGGCCAGAACTATCAAGGCTGGAGACATATCCCTGAGGAGGAGCCTGTG 842
Oy 243 AlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
|||||
Db 843 GCTTCAACCCCCACACAGTGTGATGGGGAACCAAGAGGATAAA 890

RESULT 7

US-09-904-011-206
; Sequence 206, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrata, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 206
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 973, 977, 996, 1003
; OTHER INFORMATION: unknown base
US-09-904-011-206
Alignment Scores:
Pred. No.: 1 84e-175 Length: 1620
Score: 1326.00 Matches: 257
Percent Similarity: 87.16% Conservative: 1
Best Local Similarity: 86.82% Mismatches: 0
Query Match: 97.57% Indels: 38
DB: 9 Gaps: 1
US-09-954-846-2 (1-258) x US-09-904-011-206 (1-1620)

Oy 1 MetAlaValLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrp 20
|||||
Db 3 ATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGATTCTCGTGGCGGCGACTTTCACGATG 62

OTHER INFORMATION: unknown base
US-10-175-737-31

Alignment Scores:
Pred. No.: 1,84e-175 Length: 1620
Score: 1326.00 Matches: 257
Percent Similarity: 87.16% Conservatives: 1
Best Local Similarity: 86.82% Mismatches: 0
Query Match: 97.57% Indels: 38
DB: 9 Gaps: 1

US-09-954-846-2 (1-258) x US-10-175-737-31 (1-1620)

Oy 1 MetAlaValLeuAlaProLeuIleAlaLeuValTyrSerValProHrgLeuSerArgTrp 20
Db 3 ATGGCGGTCTTGGCACCTTAATGCTCTCGTGTATTCGGTCCGCGACTTTCACGATGG 62
Oy 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuValArg 40
Db 63 CTGCGCCCAACCTTACTACCTTCTGCGGCCCTCTCTGCTGCTTCTCTACTCTCGTGAGG 122
Oy 41 LysLeuProProLeuCyshisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
Db 123 AAACGTGCGCGCTCTGCCACGCTCTGCCACCCACGAGCGGTAACCGTGTGAC 182
Oy 61 PheAspTrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsn 80
Db 183 TTTGACTGGAGAAAGTGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAAC 242
Oy 81 ArgArgSer----- 83
Db 243 CGCAGATCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCC 302
Oy 83 ----- 83
Db 303 AACCAATCTTTCTTCGCTTGGATATTCGCATGGCGCTACTTTTACATCACACTCTGC 362
Oy 84 ---MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102
Db 363 ATAGTGTCTCTGATGACGTGCAACCCGCCCTATATATGGCCCTCAGTATATCAAGTAC 422
Oy 103 PheAsnAspLysThrIleAspGluGluLeuGluArgAspLysArgValThrIleVal 122
Db 423 TTCAATGATAAACCACTGATGAGAACTAGAACGGGCAAGAGGCTCACTTGGATGTG 482
Oy 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
Db 483 GAGTTCTTTGCCAATTTGCTCTATGACTGCCAATCATTTGCCCTATCTATGCTGACCTC 542
Oy 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPhGlyLysValAspValGlyArgTyrThr 162
Db 543 TCCCTTAATACAACTGTACAGGGCTAAATTTTGGAGGTGGATTTGGACGCTATACT 602
Oy 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
Db 603 GATGTAGTACCGGTACAACTGAGCAGCATCACCCCTCACCAGCAACTCCCTACCCCTG 662
Oy 183 IleLeuPheGlnGlyLysGluAlaMetArgProGlnIleAspLysLysGlyArg 202
Db 663 ATCTGTTCACAGGTGGCAAGGCAATGCGCGCCACAGATTGCAAGAAGACCGG 722
Oy 203 AlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu 222
Db 723 GCTGTCTCATGACCTTCTCTCAGGAGAAATGTATCCGAGAAATTTAACTAAATGAGCTA 782
Oy 223 TyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVal 242
Db 783 TACCAGCGGCCCAAGAACTATCAAGGCTGGAGCAATATCCCTGAGGAGGAGCCTGTG 842
Oy 243 AlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
Db 843 GCTTCAACCCCCACCACAGTGTGATGGGGAACAAAGAGGATAAA 890

RESULT 11
US-09-906-742-206
; Sequence 206, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,742
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 206

[illegible]

```
QY 41 LysLeuProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
Db 123 AAACAGCCGCCCTCTGCCACCGGTCTGCCACCGCAACCGCAAGAGCGGTAAACCGTGTGAC 182
QY 61 PheAspTTPArgGluValGluLeuMetPheLeuSerAlaIleValMetMetLysAsn 80
Db 183 TTTGACTGGAGAGAGTGGAGATCCTGATGTTCTCAGTGCCATTGATGATGAAGAAC 242
QY 81 ArgArgSer----- 83
Db 243 CGCAGATCCATCACTGTGGAGCAACATATAGGCACACATTTTCATGTTTAGTAAAGTGCC 302
QY 83 ----- 83
Db 303 AACACAAATCTTTCTTCTCGCTTGGATATTCGCATGGGCTACTTTACATCACACTCTGC 362
QY 84 ---MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102
Db 363 ATAGTGTCTCTCATGACGTGCAACCCGCCCTATATATGGGCCCTGAGTATATCAAGTAC 422
QY 103 PheAsnAspLysThrIleAspGluGluLeuGluArgAspLysArgValThrTrpIleVal 122
Db 423 TTCATGATAAAACCATTGATGAGGAACTAGAACGGGACAAAGAGGCTCACTGGATTGTG 482
QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
Db 483 GAGTTCTTTGGCAATTTGGTCTAATGACTGCCAATCTTTGCCCTATCTATGCTGACCTC 542
QY 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr 162
Db 543 TCCTTAAATACAACTGTACAGGGCTAAATTTGGGAAGTGGATGTTGGACCTATACT 602
QY 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
Db 603 GATGTTAGTACCGGTACAAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCCCTG 662
QY 183 IleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArg 202
Db 663 ATCCCTGTTCCAAAGGTGGCAAGGAGGCAATCGCGCGGCCACAGATTGCACAAAGAGACGG 722
QY 203 AlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu 222
Db 723 GCTGTCTCATGGACCTTCTCTGAGGAGATGTGATCCGAGATTTAACTTAAATGAGCTA 782
QY 223 TyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVal 242
Db 783 TACCAGCGGGCCCAAGAAACTATCAAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTG 842
QY 243 AlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
Db 843 GUTTCACCCCCACCACAGTGTGATGGGGGAAACAAAGAGGATAAA 890
```

Search completed: July 12, 2003, 12:06:07
Job time : 142 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 09:44:17 ; Search time 201 Seconds
(without alignments)
12215.610 Million cell updates/sec

Title: US-09-954-846-4
Perfect score: 1555
Sequence: 1 AGGGAGCGGGCGGAGACC.....TAATAAAAGATTGGGATTA 1555

ring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1105431 seqs, 78947651 residues

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1555	100.0	1555	10	US-09-954-846-4
2	1511	97.2	1598	10	US-09-925-300-751
3	1381.6	88.8	1699	9	US-09-852-797-19
4	1381.6	88.8	1699	10	US-09-853-161-19
5	1381.6	88.8	1699	10	US-09-852-659A-19
6	1370	88.1	1620	9	US-09-905-291A-206
7	1370	88.1	1620	9	US-09-902-853-206
8	1370	88.1	1620	9	US-09-907-824-206
9	1370	88.1	1620	9	US-09-907-841-206
10	1370	88.1	1620	9	US-09-904-011-206
11	1370	88.1	1620	9	US-10-174-590-31
12	1370	88.1	1620	9	US-10-176-758-31
13	1370	88.1	1620	9	US-10-175-737-31
14	1370	88.1	1620	9	US-09-906-742-206
15	1370	88.1	1620	9	US-10-173-706-31
16	1370	88.1	1620	9	US-10-175-738-31
17	1370	88.1	1620	9	US-10-175-752-31
18	1370	88.1	1620	9	US-10-176-482-31
19	1370	88.1	1620	9	US-10-176-757-31

20	1370	88.1	1620	9	US-10-176-913-31	Sequence 31, Appl
21	1370	88.1	1620	9	US-10-180-552-31	Sequence 31, Appl
22	1370	88.1	1620	9	US-10-180-557-31	Sequence 31, Appl
23	1370	88.1	1620	9	US-09-906-838-206	Sequence 206, App
24	1370	88.1	1620	9	US-09-907-613-206	Sequence 206, App
25	1370	88.1	1620	9	US-09-907-942-206	Sequence 31, Appl
26	1370	88.1	1620	9	US-10-173-700-31	Sequence 31, Appl
27	1370	88.1	1620	9	US-10-174-572-31	Sequence 31, Appl
28	1370	88.1	1620	9	US-10-174-579-31	Sequence 31, Appl
29	1370	88.1	1620	9	US-10-174-582-31	Sequence 31, Appl
30	1370	88.1	1620	9	US-10-174-588-31	Sequence 31, Appl
31	1370	88.1	1620	9	US-10-175-739-31	Sequence 31, Appl
32	1370	88.1	1620	9	US-10-175-740-31	Sequence 31, Appl
33	1370	88.1	1620	9	US-10-175-743-31	Sequence 31, Appl
34	1370	88.1	1620	9	US-10-176-488-31	Sequence 31, Appl
35	1370	88.1	1620	9	US-10-176-492-31	Sequence 31, Appl
36	1370	88.1	1620	9	US-10-176-747-31	Sequence 31, Appl
37	1370	88.1	1620	9	US-10-176-750-31	Sequence 31, Appl
38	1370	88.1	1620	9	US-10-176-985-31	Sequence 31, Appl
39	1370	88.1	1620	9	US-10-176-987-31	Sequence 31, Appl
40	1370	88.1	1620	9	US-10-176-991-31	Sequence 31, Appl
41	1370	88.1	1620	9	US-10-176-992-31	Sequence 31, Appl
42	1370	88.1	1620	9	US-10-176-993-31	Sequence 31, Appl
43	1370	88.1	1620	9	US-10-184-658-31	Sequence 31, Appl
44	1370	88.1	1620	9	US-10-173-695-31	Sequence 31, Appl
45	1370	88.1	1620	9	US-10-173-697-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-954-846-4
Sequence 4, Application US/09954846
Patent No. US20020102654A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
Corley, Neil C.
Guegler, Karl J.
Baughn, Mariah R.
TITLE OF INVENTION: THIOREDOXIN PROTEINS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION NUMBER: US/09/954,846
APPLICATION DATA:
FILING DATE: 17-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/107,248
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0556 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-855-0572
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1555 base pairs

Db 26 CGGTTAGCGCGGAAAGATGGCGGCTCTGGCACTCTAATGTCTCGTATTCGTGC 85
QY 102 CCGGACTTTACAGATGGCTGCGCCAACTTACCTTCTGTGCGGCCCTCTCTCTGCTG 161
Db 86 CCGGACTTTACAGATGGCTGCGCCAACTTACCTTCTGTGCGGCCCTCTCTCTGCTG 145
QY 162 CCTTCTACTCTGAGGAACCTGCGCGCTCTGCGCACGGCTCTGCCCAACCAAGCGAAG 221
Db 146 CCTTCTACTCTGAGGAACCTGCGCGCTCTGCGCACGGCTCTGCCCAACCAAGCGAAG 205
QY 222 ACGGTAACCGCTGTGACTTTGACTGGAGAAAGTGGAGATCCTGATGTTTCTCAGTGCCA 281
Db 206 ACGGTAACCGCTGTGACTTTGACTGGAGAAAGTGGAGATCCTGATGTTTCTCAGTGCCA 265
282 TTGTGATGATGAAGAACCGCAGATCCATGTTCTCTGATGAGTGCGAAACCCCTCTATATA 341
266 TTGTGATGATGAAGAACCGCAGATCCATGTTCTCTGATGAGTGCGAAACCCCTCTATATA 325
342 TGGGCCCTGAGTATATCAAGTACTTCAATGATAAACCAATGATGAGGAACCTAGAAGCGG 401
Db 326 TGGGCCCTGAGTATATCAAGTACTTCAATGATAAACCAATGATGAGGAACCTAGAAGCGG 385
QY 402 ACAAGAGGGTCACTTGGATTGTGGAGTCTTTCGCAATTTGCTTAATGACTGCGCAATCAT 461
Db 386 ACAAGAGGGTCACTTGGATTGTGGAGTCTTTCGCAATTTGCTTAATGACTGCGCAATCAT 445
QY 462 TTGCCCTATCTATGCTGACCTCTCCCTTAAATACAACTGACAGGCTAAATTTTGGGA 521
Db 446 TTGCCCTATCTATGCTGACCTCTCCCTTAAATACAACTGACAGGCTAAATTTTGGGA 505
QY 522 AGGTGATGTTGACGCTATATGATGTAGTAGCGGGTACAAAGTGAGCACATCACCCC 581
Db 506 AGGTGATGTTGACGCTATATGATGTAGTAGCGGGTACAAAGTGAGCACATCACCCC 565
QY 582 TCACCAAGCAACTCCCTACCTCATCTGTTCAAGGTGCGCAAGGAGCAATGCGCGGC 641
Db 566 TCACCAAGCAACTCCCTACCTCATCTGTTCAAGGTGCGCAAGGAGCAATGCGCGGC 625
QY 642 CACAGATTGACAAAGAGCGGCTGTCTCATGGACCTTCTCTGAGGAGAATGTGATCC 701
Db 626 CACAGATTGACAAAGAGCGGCTGTCTCATGGACCTTCTCTGAGGAGAATGTGATCC 685
702 GAGAATTTAACTTAATGAGCTATACAGCGGCGCAAGAACTATCAAGGCTGGAGACA 761
686 GAGAATTTAACTTAATGAGCTATACAGCGGCGCAAGAACTATCAAGGCTGGAGACA 745
QY 762 ATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACAGTGTGAGATGGGGAACA 821
Db 746 ATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACAGTGTGAGATGGGGAACA 805
QY 822 AGAAGGATAAATAAGATCCTCACTTTGGCAGTGTCTCTCTCTGTCCTCAATTCAGGCTCT 881
Db 806 AGAAGGATAAATAAGATCCTCACTTTGGCAGTGTCTCTCTCTGTCCTCAATTCAGGCTCT 865
QY 882 TTCCATAACACAGGCTGAGGCTGAGGCTTTTATTTATGTTTTCCTTTGCGTGTGAC 941
Db 866 TTCCATAACACAGGCTGAGGCTGAGGCTTTTATTTATGTTTTCCTTTGCGTGTGAC 925
QY 942 TGGGTGGGCGAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATGTGAGGCACC 1001
Db 926 TGGGTGGGCGAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATGTGAGGCACC 985
QY 1002 CTACAGGAAGGCTGCCATGCTGTGGCCAACTGTTTTCACCTGGAGCAAGAAAGAGATCTCA 1061
Db 986 CTACAGGAAGGCTGCCATGCTGTGGCCAACTGTTTTCACCTGGAGCAAGAAAGAGATCTCA 1045
QY 1062 TAGGACGGAGGGGAAATGGTTTCCCTTCCAAAGCTTGGGTGAGTGTGTTAACTGCTTATCA 1121
Db 1046 TAGGACGGAGGGGAAATGGTTTCCCTTCCAAAGCTTGGGTGAGTGTGTTAACTGCTTATCA 1105
QY 1122 GCTATTTCAGACATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATCTCTCTTA 1181

Db 1106 GCTATTTCAGACATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATCTCTCTTA 1165
QY 1182 GTTGACCTTCACAGCTTGGTTAGACCTTAGATTTAAACCTTAAGSTAAGATGCTGGGGTATA 1241
Db 1166 GTTGACCTTCACAGCTTGGTTAGACCTTAGATTTAAACCTTAAGSTAAGATGCTGGGGTATA 1225
QY 1242 GAACGCTAAAGAAATTTTCCCAAGGACTCTTGTCTTCCCTTAAGCCCTTCTTGCTTCGTTTA 1301
Db 1226 GAACGCTAAAGAAATTTTCCCAAGGACTCTTGTCTTCCCTTAAGCCCTTCTTGCTTCGTTTA 1285
QY 1302 TGCTTCTTCAATTAAGATATAAGCCTAACTTTGTGCTAGTCTTAAAGAGAAACCTTTAAC 1361
Db 1286 TGCTTCTTCAATTAAGATATAAGCCTAACTTTGTGCTAGTCTTAAAGAGAAACCTTTAAC 1345
QY 1362 CACAAAGTTTTATCATTTGAAGACAATTTGAACAACCCCTATTTCTGGGGATTGAGA 1421
Db 1346 CACAAAGTTTTATCATTTGAAGACAATTTGAACAACCCCTATTTCTGGGGATTGAGA 1405
QY 1422 AGGGTCAATAGAGGCTTGAGACTTTCTTGTGTGTAGGACTTGGAGGAGAAATCCCC 1481
Db 1406 AGGGTCAATAGAGGCTTGAGACTTTCTTGTGTGTAGGACTTGGAGGAGAAATCCCC 1465
QY 1482 TGGACTTTCATAACCTCTGACATACCTCCCAACCCAGTGTGATGCTTTCCGTAAATAA 1541
Db 1466 TGGACTTTCATAACCTCTGACATACCTCCCAACCCAGTGTGATGCTTTCCGTAAATAA 1525
QY 1542 AAAGATTGGGATT 1554
Db 1526 AAAGATTGGGATT 1538

RESULT 3

US-09-852-797-19
; Sequence 19, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 19:
; LENGTH: 1699
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE

LOCATION: (871)
OTHER INFORMATION: n equals a,t,g, or c
US-09-852-797-19

Query Match 88.8%; Score 1381.6; DB 9; Length 1699;

Best Local Similarity 92.3%; Pred. No. 0;

Matches 1511; Conservative 1; Mismatches 11; Indels 114; Gaps 1;

QY	32	CGAGCAGTGGCGGTTACGGCCCAAGAGATGGCGGCTTGGCACCTCTAATTGCTCTCGTG	91
DB	3	CGTGGCGAATTGGGCACAGCGCAAGAGATGGCGGCTTGGCACCTCTAATTGCTCTCGTG	62
QY	92	TATTCGGTGGCGCGACTTTCAGATGGCTGCGCCCAACCTTACTACCTCTCTGCGGCCCTG	151
DB	63	TATTCGGTGGCGCGACTTTCAGATGGCTGCGCCCAACCTTACTACCTCTCTGCGGCCCTG	122
QY	152	CTCTCTGCTGCTTCTCTACTCTGTGAGGAAACTGCGCGCGCTCTGCCACGGTGTGCCAC	211
DB	123	CTCTCTGCTGCTTCTCTACTCTGTGAGGAAACTGCGCGCGCTCTGCCACGGTGTGCCAC	182
QY	212	CAACCGAAGAGGTTAACCCTGTGACTTGTGACTGGAGAGAAGTGGAGATCTCTGATGTTT	271
DB	183	CAACCGAAGAGGTTAACCCTGTGACTTGTGACTGGAGAGAAGTGGAGATCTCTGATGTTT	242
QY	272	CTCAGTGCCATTGTGATGATGAAGAACCGCAGATCCA	308
DB	243	CTCAGTGCCATTGTGATGATGAAGAACCGCAGATCCA	302
QY	309	-----	308
DB	303	AACATTTTCATGTTAGTAAAGTGCCCAACACAATTTCTTCTTCGGCTTGGATATTCG	362
QY	309	-----TGTTCTCTGATGACGTGCAAAACCCCTTA	337
DB	363	ATGGGCCCTACTTACATCACACTCTGCATAGTGTTCCTGATGACGTGCAAAACCCCTTA	422
QY	338	TATATGGGCCCTGAGTATATCAAGTACTTCAATGATATAAACCAATTCATGAGGAACCTAGAA	397
DB	423	TATATGGGCCCTGAGTATATCAAGTACTTCAATGATATAAACCAATTCATGAGGAACCTAGAA	482
QY	398	CGGACACAGAGGTCACCTGGATTTGTGGATTTCTTTGGCAATTTGGTCTAATGACTGCCAA	457
DB	483	CGGACACAGAGGTCACCTGGATTTGTGGATTTCTTTGGCAATTTGGTCTAATGACTGCCAA	542
QY	458	TCATTTGGCCCTATCTATGCTGACCTCTCCCTTAATACAACTGTACAGGGCTTAATTTT	517
DB	543	TCATTTGGCCCTATCTATGCTGACCTCTCCCTTAATACAACTGTACAGGGCTTAATTTT	602
QY	518	GGGAGGTGGATTTGGACGCTATATCTGATGTTAGTACGCTGACAAAGTGAGCATCA	577
DB	603	GGGAGGTGGATTTGGACGCTATATCTGATGTTAGTACGCTGACAAAGTGAGCATCA	562
QY	578	CCCTTCACCAAGCACTCCCTACCTGATCTCTGTTCCAAAGTGGCAAGGAGGCAATGCGG	637
DB	663	CCCTTCACCAAGCACTCCCTACCTGATCTCTGTTCCAAAGTGGCAAGGAGGCAATGCGG	722
QY	638	CGGCCACAGATTGACAAGAGGAGCGGCTGTCTCATGACCTTCTCTGAGGAGATGTG	697
DB	723	CGGCCACAGATTGACAAGAGGAGCGGCTGTCTCATGACCTTCTCTGAGGAGATGTG	782
QY	698	ATCCGAGAAATTAACCTTAATGAGCTATACCAAGCGGGCCAAAGAACTATCAAGGCTGGA	757
DB	783	ATCCGAGAAATTAACCTTAATGAGCTATACCAAGCGGGCCAAAGAACTATCAAGGCTGGA	842
QY	758	GACATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCGCACACAGTGTCTGATGGGAA	817
DB	843	GACATATCCCTGAGGAGCAGCCTGTGGNTTCAACCCCGCACACAGTGTCTGATGGGAA	902
QY	818	AACAGAGGATAAATAGATCTTCACTTGGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT	877
DB	903	AACAGAGGATAAATAGATCTTCACTTGGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT	962
QY	878	CTCTTTCCATAACCAACAGCCTGAGGCTGCAGCCCTTTTATTTATGTTTTTCCCTTTGGCTG	937

DB	963	CTCTTTCCATAACCAACAGCCTGAGGCTGAGCCCTTTATTTATGTTTTTCCCTTTGGCTG	1022
QY	938	TGACTGGGTGGGCGAGCATGCGAGCTTCTGATTTTAAAGAGGCATCTAGGGAATGTCAGG	997
DB	1023	TGACTGGGTGGGCGAGCATGCGAGCTTCTGATTTTAAAGAGGCATCTAGGGAATGTCAGG	1082
QY	998	CACCCTACAGGAGGCGCTGCCATGCTGTGCCCAACTGTTTCACTGGAGCAAGAAAGAT	1057
DB	1083	CACCCTACAGGAGGCGCTGCCATGCTGTGCCCAACTGTTTCACTGGAGCAAGAAAGAT	1142
QY	1058	CTCATAGGAGGAGGGGAAATGGTTTCCCTTCAAGCTTGGGTGAGTGTGTTAACTGCTT	1117
DB	1143	CTCATAGGAGGAGGGGAAATGGTTTCCCTTCAAGCTTGGGTGAGTGTGTTAACTGCTT	1202
QY	1118	ATCAGCTATTTCAGACATCTCCATGCTTCTCCATGAACCTCTGCTGTTTTCATCTCCTT	1177
DB	1203	ATCAGCTATTTCAGACATCTCCATGCTTCTCCATGAACCTCTGCTGTTTTCATCTCCTT	1262
QY	1178	CTTAGTTGACCTGCGACAGCTTGGTTAGACCTAGATTTAACCTTAAGCTAAGATGCTGGG	1237
DB	1263	CTTAGTTGACCTGCGACAGCTTGGTTAGACCTAGATTTAACCTTAAGCTAAGATGCTGGG	1322
QY	1238	TATAGAACCTAAGAAATTTTCCCCCAAGGACTCTTCTCTTCCCTTAAGCCCTTCTGGCTTCG	1297
DB	1323	TATAGAACCTAAGAAATTTTCCCCCAAGGACTCTTCTCTTCCCTTAAGCCCTTCTGGCTTCG	1382
QY	1298	TTTATGGCTTCTCATTTAAAGTATAAGCCCTTAACCTTCTGCTAGTCTCTAAGGAGAAACCTT	1357
DB	1383	TTTATGGCTTCTCATTTAAAGTATAAGCCCTTAACCTTCTGCTAGTCTCTAAGGAGAAACCTT	1442
QY	1358	TAAACCAAAAGTTTATCATTTCAAGACAATATTGAACAAACCCCTATTTTGTGGGAT	1417
DB	1443	TAAACCAAAAGTTTATCATTTCAAGACAATATTGAACAAACCCCTATTTTGTGGGAT	1502
QY	1418	GAGAAGGGTGAATAGAGCTTTCAGACTTTCCTTTTGTGGTAGGACTTGGAGGAGAAAT	1477
DB	1503	GAGAAGGGTGAATAGAGCTTTCAGACTTTCCTTTTGTGGTAGGACTTGGAGGAGAAAT	1562
QY	1478	CCCTTGGACTTTCACCTAACCTTCTGACATACCTCCACACCCAGCTTGTATGGCTTCCGTA	1537
DB	1563	CCCTTGGACTTTCACCTAACCTTCTGACATACCTCCACACCCAGCTTGTATGGCTTCCGTA	1622
QY	1538	ATAAAAAGATTGGGAT	1554
DB	1623	ATAAAAAGATTGGGAT	1639

RESULT 4

US-09-853-161-19
Sequence 19, Application US/09853161
Patent No. US20020076756A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 28 Human Secreted Proteins
FILE REFERENCE: P2003P3
CURRENT APPLICATION NUMBER: US/09/853,161
CURRENT FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: 60/265,583
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 09/152,060
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: PCT/US98/04858
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/040,762
PRIOR FILING DATE: 1997-03-14
PRIOR APPLICATION NUMBER: 60/040,710
PRIOR FILING DATE: 1997-03-14
PRIOR APPLICATION NUMBER: 60/050,934
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: 60/048,100
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: 60/048,357

QY	638	CGGCCACAGATTGACAGAAAGGACGGGCTGTCTCTATGAGACCTTCTCTGAGGAGAAATGTG	697
DB	723	CGGCCACAGATTGACAGAAAGGACGGGCTGTCTCTATGAGACCTTCTCTGAGGAGAAATGTG	782
QY	698	ATCCGAGAAATTAACTTAAATAGACCTATACACAGCGGCCAAGAACAATCATCAAAAGGCTGGA	757
DB	793	ATCCGAGAAATTAACTTAAATAGACCTATACACAGCGGCCAAGAACAATCATCAAAAGGCTGGA	842
QY	758	GACAAATATCCCTGAGGAGCAGCCTGTGGCTTTCAACCCCCCACCACAGAGTGTCCAGATGGGAA	817
DB	843	GACAAATATCCCTGAGGAGCAGCCTGTGGCTTTCAACCCCCCACCACAGAGTGTCCAGATGGGAA	902
QY	818	AACAGAAAGGATAAATAGATCTCTACTTTGGGAGTGGTTCCTCTCTCTGCAATTCAGG	877
DB	903	AACAGAAAGGATAAATAGATCTCTACTTTGGGAGTGGTTCCTCTCTCTGCAATTCAGG	962
QY	878	CTCTTTCCATAACCCACAGCCTGAGGCTGACGCCCTTTATTATATGTTTTCCCTTTGGCGTG	937
DB	963	CTCTTTCCATAACCCACAGCCTGAGGCTGACGCCCTTTATTATATGTTTTCCCTTTGGCGTG	1022
QY	938	TGACTTGGTGGGGCAGCATGCAGCTTCTGATTTTTAAAGAGGSCATCTAGGGAATTTGTCAGG	997
DB	1023	TGACTTGGTGGGGCAGCATGCAGCTTCTGATTTTTAAAGAGGSCATCTAGGGAATTTGTCAGG	1082
QY	998	CACCGTACAGAAAGCCTGCCATGCTGTGGCCAACTGTTTCACATGGGAGCAGAAGAAGAT	1057
DB	1083	CACCGTACAGAAAGCCTGCCATGCTGTGGCCAACTGTTTCACATGGGAGCAGAAGAAGAT	1142
QY	1058	CTCATAGACGAGGGGGAAATGTTTCCCTCCCAAGCTTGGGTCAGAGTGTGTTAACTGCTT	1117
DB	1143	CTCATAGACGAGGGGGAAATGTTTCCCTCCCAAGCTTGGGTCAGAGTGTGTTAACTGCTT	1202
QY	1118	ATCAGCTATTACAGACATCTCCATGGTTTCTCCATGAAACTGTGTGGTTTCATCATTCCTT	1177
DB	1203	ATCAGCTATTACAGACATCTCCATGGTTTCTCCATGAAACTGTGTGGTTTCATCATTCCTT	1262
QY	1178	CTTAGTTGACCTGCACAGCTTGGTTAGACCTAGATTTAAACCTTAAGTATAGATGCTGGG	1237
DB	1263	CTTAGTTGACCTGCACAGCTTGGTTAGACCTAGATTTAAACCTTAAGTATAGATGCTGGG	1322
QY	1238	TATAGAAGCGCTAAGAAATTTTCCCCCAAGGACTCTTGCTTCTTAAAGCCCTTCTGGCTTCG	1297
DB	1323	TATAGAAGCGCTAAGAAATTTTCCCCCAAGGACTCTTGCTTCTTAAAGCCCTTCTGGCTTCG	1382
QY	1298	TTTATGGTCTTCATTTAAAGTATTAAGCCTAACTTTTGTGCGTAGTTCCTTAAGGAGAAACCTT	1357
DB	1383	TTTATGGTCTTCATTTAAAGTATTAAGCCTAACTTTTGTGCGTAGTTCCTTAAGGAGAAACCTT	1442
QY	1358	TAACCCACAAGCTTTTATCATTTGAAGACAATATTGAACACCCCTATTTTGTGGGGATT	1417
DB	1443	TARCCACAAGCTTTTATCATTTGAAGACAATATTGAACACCCCTATTTTGTGGGGATT	1502
QY	1418	GAGAAGGGGTGAATAGAGGCTTGAGACTTTCCCTTTGTGGTAGGACTTGGAGGAGAAAT	1477
DB	1503	GAGAAGGGGTGAATAGAGGCTTGAGACTTTCCCTTTGTGGTAGGACTTGGAGGAGAAAT	1562
QY	1478	CCCTGGACTTTTCACTAACCCCTTGACATACTCCCCACACCCAGTTGATGGCTTTCCGTA	1537
DB	1563	CCCTGGACTTTTCACTAACCCCTTGACATACTCCCCACACCCAGTTGATGGCTTTCCGTA	1622
QY	1538	ATAAAAAGATTGGGATT	1554
DB	1623	ATRAAAAAGATTGGGATT	1639

RESULT 5

US-09-852-659A-19

; Sequence 19, Application US/09852659A

; Patent No. US20020077287A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 28 Human Secreted Proteins

;; CURRENT APPLICATION NUMBER: US/09/852,659A

;; PRIOR FILING DATE: 2001-05-11

;; PRIOR APPLICATION NUMBER: 60/265,583

;; PRIOR FILING DATE: 2001-02-02

;; PRIOR APPLICATION NUMBER: 09/152,060

;; PRIOR FILING DATE: 1998-09-11

;; PRIOR APPLICATION NUMBER: PCT/US98/04858

;; PRIOR FILING DATE: 1998-03-12

;; PRIOR APPLICATION NUMBER: 60/040,762

;; PRIOR FILING DATE: 1997-03-14

;; PRIOR APPLICATION NUMBER: 60/040,710

;; PRIOR FILING DATE: 1997-03-14

;; PRIOR APPLICATION NUMBER: 60/050,934

;; PRIOR FILING DATE: 1997-05-30

;; PRIOR APPLICATION NUMBER: 60/048,100

;; PRIOR FILING DATE: 1997-05-30

;; PRIOR APPLICATION NUMBER: 60/048,357

;; PRIOR FILING DATE: 1997-05-30

;; PRIOR APPLICATION NUMBER: 60/048,189

;; PRIOR FILING DATE: 1997-05-30

;; PRIOR APPLICATION NUMBER: 60/057,765

;; PRIOR FILING DATE: 1997-09-05

;; PRIOR APPLICATION NUMBER: 60/048,970

;; PRIOR FILING DATE: 1997-06-06

;; PRIOR APPLICATION NUMBER: 60/068,368

;; PRIOR FILING DATE: 1997-12-19

;; NUMBER OF SEQ ID NOS: 121

;; SOFTWARE: PatentIn ver. 2.0

;; SEQ ID NO 19

;; LENGTH: 1699

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: SITE

;; LOCATION: (871)

;; OTHER INFORMATION: n equals a,t,g, or c

US-09-852-659A-19

Query Match 88.8%; Score 1381.6; DB 10; Length 1699;

Best Local Similarity 92.3%; Pred. No. 0;

Matches 1511; Conservative 1; Mismatches 11; Indels 114; Gaps 1;

QY	32	CGACGAGGCGGTTACGCCGCAAGAGATGGCGGTTTGGCACTTCTAATTCCTCGTG	91
DB	1		
QY	3	CGTCCGAAATTCGCACGAGCGAAAGATGGCGGTTTGGCACTTCTAATTCCTCGTG	62
DB	1		
QY	92	TATTCGGTGGCGGACTTTCACGATGGCTCGGCCCAACCTTACTACCTTCTGTGCGCCCTG	151
DB	1		
QY	63	TATTCGGTGGCGGACTTTCACGATGGCTCGGCCCAACCTTACTACCTTCTGTGCGCCCTG	122
DB	1		
QY	152	CTCTCTGCTGCTTCTCTACTCGTGAGGAACTGCGCGGCTGCGCAGGCTGCGCCACC	211
DB	1		
QY	123	CTCTCTGCTGCTTCTCTACTCGTGAGGAACTGCGCGGCTGCGCAGGCTGCGCCACC	182
DB	1		
QY	212	CAACCGAAGACGGTAACCGCTGTGACTTTGACTGGAGAGAGTGGAGATCTGTATGTTT	271
DB	1		
QY	183	CAACCGAAGACGGTAACCGCTGTGACTTTGACTGGAGAGAGTGGAGATCTGTATGTTT	242
DB	1		
QY	272	CTCAGTGCATTTGATGATGAAGAACCGCAGATCCA	308
DB	1		
QY	243	CTCAGTGCATTTGATGATGAAGAACCGCAGATCCA	302
DB	1		
QY	309	-----	308
DB	1	-----	
QY	303	AACATTTTCATGTTTAGTAAAGTGCCCAACAAATCTTTTCTCCGCTTGGATATTCGC	362
DB	1		
QY	309	-----TGTTCTGATGACGTGCAAAACCCCTTA	337
DB	1		
QY	363	ATGGCCCTACTTTACATCACACTCTGCATAGTGTCTCTGATGACGTGCAAAACCCCTTA	422
DB	1		
QY	338	TATATGGCCCTGAGTATATCAAGTACTTCAATGATAAAACCATTCATGAGGAAGTAA	397
DB	1		
QY	423	TATATGGCCCTGAGTATATCAAGTACTTCAATGATAAAACCATTCATGAGGAAGTAA	482
DB	1		

QY	398	CGGGACAAGAGGTCACCTGGGATTTGGAGTTCTTTGCCAATTTGGTCTAATGACTGCCAA	457
DB	1		
QY	483	CGGGACAAGAGGTCACCTGGGATTTGGAGTTCTTTGCCAATTTGGTCTAATGACTGCCAA	542
DB	1		
QY	458	TCATTTGCCCTTATCTATGCTGACCTCTCCCTTAATACAACTGTACAGGGCTAAATTTT	517
DB	1		
QY	543	TCATTTGCCCTTATCTATGCTGACCTCTCCCTTAATACAACTGTACAGGGCTAAATTTT	602
DB	1		
QY	518	GGGAAGGTGGATTTGGAGGCTATCTATGATTTAGTACGCGGTACAAAGTACAGCATCA	577
DB	1		
QY	603	GGGAAGGTGGATTTGGAGGCTATCTATGATTTAGTACGCGGTACAAAGTACAGCATCA	662
DB	1		
QY	578	CCCCCACCACCAACTCCCTACCTGATCCTGTTTCCAAAGGTGGCAAGGCAATGCCG	637
DB	1		
QY	663	CCCCCACCACCAACTCCCTACCTGATCCTGTTTCCAAAGGTGGCAAGGCAATGCCG	722
DB	1		
QY	638	CGGCCACAGATTGACAAGAAAGGAGCGGCTGTCTCATGGACCTTCTCTGAGGAGATGTG	697
DB	1		
QY	723	CGGCCACAGATTGACAAGAAAGGAGCGGCTGTCTCATGGACCTTCTCTGAGGAGATGTG	782
DB	1		
QY	698	ATCCGAGAAATTTAACTTAAATGAGCTATACAGCGGGCCAAAGAACTATCAAGGCTGGA	757
DB	1		
QY	783	ATCCGAGAAATTTAACTTAAATGAGCTATACAGCGGGCCAAAGAACTATCAAGGCTGGA	842
DB	1		
QY	758	GACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCCACACACAGTGTACAGATGGGA	817
DB	1		
QY	843	GACAATATCCCTGAGGAGCAGCTGTGGNTTCAACCCCCACACACAGTGTACAGATGGGA	902
DB	1		
QY	818	AACAAGAGGATAAATAAGATCCTCCTTTGGCAGTGTCTTCTCTCTCTCTCTCAATTCAGG	877
DB	1		
QY	903	AACAAGAGGATAAATAAGATCCTCCTTTGGCAGTGTCTTCTCTCTCTCTCAATTCAGG	962
DB	1		
QY	878	CTCTTTCCATAACCAAGCCCTGAGGCTGACGCTTTTATTTATTTATTTATTTATTTATTT	937
DB	1		
QY	963	CTCTTTCCATAACCAAGCCCTGAGGCTGACGCTTTTATTTATTTATTTATTTATTTATTT	1022
DB	1		
QY	938	TCAGTGGGTGGGCGACATGACCTTCTGATTTTAAAGAGGATCTAGGAAATGTCAGG	997
DB	1		
QY	1023	TCAGTGGGTGGGCGACATGACCTTCTGATTTTAAAGAGGATCTAGGAAATGTCAGG	1082
DB	1		
QY	998	CACCTTACAGGAGGCTGCCATGCTGTGGCCAACTGTTTCTACTGGAGCAAGAAAGAGAT	1057
DB	1		
QY	1083	CACCTTACAGGAGGCTGCCATGCTGTGGCCAACTGTTTCTACTGGAGCAAGAAAGAGAT	1142
DB	1		
QY	1058	CTCATAGGAGGAGGGGAAATGGTTTCCCTCCCAAGCTTGGGTGAGTGTGTTAACTGCTT	1117
DB	1		
QY	1143	CTCATAGGAGGAGGGGAAATGGTTTCCCTCCCAAGCTTGGGTGAGTGTGTTAACTGCTT	1202
DB	1		
QY	1118	ATCAGCTATTACAGATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATTCCTT	1177
DB	1		
QY	1203	ATCAGCTATTACAGATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATTCCTT	1262
DB	1		
QY	1178	CTTAGTTGACCTGCAGCTTGGTTAGACCTAGATTAAACCCCTAAGGTAAAGTGTGGG	1237
DB	1		
QY	1263	CTTAGTTGACCTGCAGCTTGGTTAGACCTAGATTAAACCCCTAAGGTAAAGTGTGGG	1322
DB	1		
QY	1238	TATAGACGCTAAGAAATTTTCCCAAGGACTCTTGTCTTCCCTTAAGCCCTTCTGGCTTCG	1297
DB	1		
QY	1323	TATAGACGCTAAGAAATTTTCCCAAGGACTCTTGTCTTCCCTTAAGCCCTTCTGGCTTCG	1382
DB	1		
QY	1298	TTTATGGTCTTCAATAAGATATAAGCCTTAACCTTTGTCGCTAGTCTCTAGGAGAACTT	1357
DB	1		
QY	1383	TTTATGGTCTTCAATAAGATATAAGCCTTAACCTTTGTCGCTAGTCTCTAGGAGAACTT	1442
DB	1		
QY	1358	TAAACACAAAGTTTTTATCATTGAAGACAATATTGAACAAACCCCTTATTTTGTGGGATT	1417
DB	1		
QY	1443	TAAACACAAAGTTTTTATCATTGAAGACAATATTGAACAAACCCCTTATTTTGTGGGATT	1502
DB	1		
QY	1418	GAGAAGGGTGAATAGAGCTTTCAGACTTTCCTTTTGTGTGTAGGACTTTGGAGGAGAAAT	1477
DB	1		
QY	1503	GAGAAGGGTGAATAGAGCTTTCAGACTTTCCTTTTGTGTGTAGGACTTTGGAGGAGAAAT	1562
DB	1		

QY 1478 CCCTGGACTTTCACCTAACCTCTGACATACCTCCACACACCCAGTTGATGGCTTTCGGTA 1537
|||||
Db 1563 CCCTGGACTTTCACCTAACCTCTGACATACCTCCACACACCCAGTTGATGGCTTTCGGTA 1622
|||||
QY 1538 ATAAAAAGATTGGATT 1554
|||||
Db 1623 ATAAAAAGATTGGATT 1639
|||||

RESULT 6

US-09-905-291A-206

: Sequence 206, Application US/09905291A

: Patent No. US20020160374A1

: GENERAL INFORMATION:

: APPLICANT: Genentech, Inc.

: APPLICANT: Ashkenazi, Avi

: APPLICANT: Botstein, David

: APPLICANT: Desnoyers, Luc

: APPLICANT: Eaton, Dan L.

: APPLICANT: Ferrara, Napoleone

: APPLICANT: Filvaroff, Ellen

: APPLICANT: Fong, Sherman

: APPLICANT: Gao, Wei-Qiang

: APPLICANT: Gerber, Hanspeter

: APPLICANT: Gerritsen, Mary E.

: APPLICANT: Goddard, A.

: APPLICANT: Godowski, Paul J.

: APPLICANT: Grimaldi, Christopher J.

: APPLICANT: Gurney, Austin L.

: APPLICANT: Hillan, Kenneth, J.

: APPLICANT: Kijavini, Ivar J.

: APPLICANT: Mather, Jennie P.

: APPLICANT: Pan, James

: APPLICANT: Paoni, Nicholas F.

: APPLICANT: Roy, Margaret Ann

: APPLICANT: Stewart, Timothy A.

: APPLICANT: Tumas, Daniel

: APPLICANT: Williams, P. Mickey

: APPLICANT: Wood, William, I.

: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

: TITLE OF INVENTION: Acids Encoding the Same

: FILE REFERENCE: 10466-14

: CURRENT APPLICATION NUMBER: US/09/905,291A

: PRIOR FILING DATE: 2001-07-12

: PRIOR APPLICATION NUMBER: PCT/US00/04414

: PRIOR FILING DATE: 2000-02-22

: PRIOR APPLICATION NUMBER: US 60/143,048

: PRIOR FILING DATE: 1999-07-07

: PRIOR APPLICATION NUMBER: US 60/145,698

: PRIOR FILING DATE: 1999-07-26

: PRIOR APPLICATION NUMBER: US 60/146,222

: PRIOR FILING DATE: 1999-07-28

: PRIOR APPLICATION NUMBER: PCT/US99/20594

: PRIOR FILING DATE: 1999-09-08

: PRIOR APPLICATION NUMBER: PCT/US99/20944

: PRIOR FILING DATE: 1999-09-13

: PRIOR APPLICATION NUMBER: PCT/US99/21090

: PRIOR FILING DATE: 1999-09-15

: PRIOR APPLICATION NUMBER: PCT/US99/21547

: PRIOR FILING DATE: 1999-09-15

: PRIOR APPLICATION NUMBER: PCT/US99/23089

: PRIOR FILING DATE: 1999-10-05

: PRIOR APPLICATION NUMBER: PCT/US99/28214

: PRIOR FILING DATE: 1999-11-29

: PRIOR APPLICATION NUMBER: PCT/US99/28313

: PRIOR FILING DATE: 1999-11-30

: PRIOR APPLICATION NUMBER: PCT/US99/28564

: PRIOR FILING DATE: 1999-12-02

: PRIOR APPLICATION NUMBER: PCT/US99/28565

: PRIOR FILING DATE: 1999-12-02

: PRIOR APPLICATION NUMBER: PCT/US99/30095

: PRIOR FILING DATE: 1999-12-16

: PRIOR APPLICATION NUMBER: PCT/US99/30911

: PRIOR FILING DATE: 1999-12-16

: PRIOR APPLICATION NUMBER: PCT/US99/30911

: PRIOR FILING DATE: 1999-12-16

: PRIOR APPLICATION NUMBER: PCT/US99/30911

: PRIOR FILING DATE: 1999-12-16

: PRIOR APPLICATION NUMBER: PCT/US99/30911

: PRIOR FILING DATE: 1999-12-16

: PRIOR APPLICATION NUMBER: PCT/US99/30911

: PRIOR FILING DATE: 1999-12-16

: PRIOR FILING DATE: 1999-12-20
: PRIOR APPLICATION NUMBER: PCT/US99/30999
: PRIOR FILING DATE: 1999-12-20
: PRIOR APPLICATION NUMBER: PCT/US00/00219
: PRIOR FILING DATE: 2000-01-05
: NUMBER OF SEQ ID NOS: 423
: SEQ ID NO 206
: LENGTH: 1620
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (973)
: OTHER INFORMATION: a, t, c or g
: NAME/KEY: modified_base
: LOCATION: (977)
: OTHER INFORMATION: a, t, c or g
: NAME/KEY: modified_base
: LOCATION: (996)
: OTHER INFORMATION: a, t, c or g
: NAME/KEY: modified_base
: LOCATION: (1003)
: OTHER INFORMATION: a, t, c or g
: US-09-905-291A-206

Query Match 88.1%; Score 1370; DB 9; Length 1620;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1494; Conservative 0; Mismatches 4; Indels 114; Gaps 1;

QY 57 AGATGGCGGTCTTGGCACCTCTAATTCCTCTCGTGTATTTCGTCGCGGACCTTTCCAGAT 116
|||||
Db 1 AGATGGCGGTCTTGGCACCTCTAATTCCTCTCGTGTATTTCGTCGCGGACCTTTCCAGAT 60
|||||
QY 117 GGCTCGGCCAACCTTACTACCTTCTGTGCGGCCCTGCTCTCTGCTGCTTCTCTACTCGTGA 176
|||||
Db 61 GGCTCGGCCAACCTTACTACCTTCTGTGCGGCCCTGCTCTCTGCTGCTTCTCTACTCGTGA 120
|||||
QY 177 GGAACCTGCGCGCTCTGCGCGGTCTGCCACCCCAACCGGAGAGCGGTAAACCGGTGTG 236
|||||
Db 121 GGAACCTGCGCGCTCTGCGCGGTCTGCCACCCCAACCGGAGAGCGGTAAACCGGTGTG 180
|||||
QY 237 ACTTTGACTGGAGAGAGTGGAGATCTGATGTTTCTCAGTGGCCATGTGTATGATGAAGA 296
|||||
Db 181 ACTTTGACTGGAGAGAGTGGAGATCTGATGTTTCTCAGTGGCCATGTGTATGATGAAGA 240
|||||
QY 297 ACCGCAGATCCA----- 308
|||||
Db 241 ACCGCAGATCCA----- 300
|||||
QY 309----- 308
|||||
Db 301 CCAACACAATTTCTTCTCCGCTTGGATATTTCGATGGGCTACTTTACATCACACTCT 360
|||||
QY 309-----TGTTCCTGATGACGTGCAACCCGCCCTATATATGCGGCCCTGAGTATATCAAGT 362
|||||
Db 361 GCATAGTGTCTCTGATGACGTGCAACCCGCCCTATATATGCGGCCCTGAGTATATCAAGT 420
|||||
QY 363 ACTTCAATGATAAAACCATTTGATGAGGAACCTAGAACGGGACCAAGAGGTCTACTTGGATTG 422
|||||
Db 421 ACTTCAATGATAAAACCATTTGATGAGGAACCTAGAACGGGACCAAGAGGTCTACTTGGATTG 480
|||||
QY 423 TGGAGTTCTTTTGGCAATTTGGTCTAATGACTGCCAATCATTTGCCCTCTATCTATCTGACC 482
|||||
Db 481 TGGAGTTCTTTTGGCAATTTGGTCTAATGACTGCCAATCATTTGCCCTCTATCTATCTGACC 540
|||||
QY 483 TCTCCCTTAATAACACTGTACAGGGCTAAATTTGGGAAGGTGGATGTTGGAGCGCTATA 542
|||||
Db 541 TCTCCCTTAATAACACTGTACAGGGCTAAATTTGGGAAGGTGGATGTTGGAGCGCTATA 600
|||||
QY 543 CTGATGTTAGTACCGGTACAAAAGTACACATCACCCTCACCACCAAGCACTCCCTACCC 602
|||||
Db 601 CTGATGTTAGTACCGGTACAAAAGTACACATCACCCTCACCACCAAGCACTCCCTACCC 660
|||||

Qy	603	TGATCTCTGTTCCAAAGGTGGCAAGAGAGGCAATGCGCGCGCCACACAGATTGACAGAAAGGAC	662
Db	661	TGATCTCTGTTCCAAAGGTGGCAAGAGAGCAATGCGCGCGCCACAGATTGACAGAAAGGAC	720
Qy	663	GGGCTGTCATAGCACCTTCTCTCAGGAGAAATGTATCCGAGAAATTTAACTTTAAATGAGC	722
Db	721	GGGCTGTCATAGCACCTTCTCTCAGGAGAAATGTATCCGAGAAATTTAACTTTAAATGAGC	780
Qy	723	TATACCAAGCGGGCCCAAGAAACTATCAAGGCTGGAGACAATATCCCTTGAGGAGCAGCCTG	782
Db	781	TATACCAAGCGGGCCCAAGAAACTATCAAGGCTGGAGACAATATCCCTTGAGGAGCAGCCTG	840
Qy	783	TGGCTTCAACCCCAACACACAGTGTACATGGGGAAAAACAAGAGGATAAATAAGATCCTC	842
Db	841	TGGCTTCAACCCCAACACACAGTGTACATGGGGAAAAACAAGAGGATAAATAAGATCCTC	900
Qy	843	ACTTTGGCAGTGTCTCTCTCTCAATGCCAGGCTTTTCCATAACCAACAAGCCCTGAG	902
Db	901	ACTTTGGCAGTGTCTCTCTCTCAATGCCAGGCTTTTCCATAACCAACAAGCCCTGAG	960
Qy	903	GCTGCAGCCCTTTTATTTATGTTTTCCTTTGGCTGTGACTGGGTGGGCAAGCATGCAAGCT	962
Db	961	GCTGCAGCCCTTTTATTTATGTTTTCCTTTGGCTGTGACTGGGTGGGCAAGCATGCAAGCT	1020
	963	TCTGATTTTAAAGAGGCATCTAGGGAATTTGTTCAGGCACCCCTACAGGAAGGCGTGCCTATGC	1022
Db	1021	TCTGATTTTAAAGAGGCATCTAGGGAATTTGTTCAGGCACCCCTACAGGAAGGCGTGCCTATGC	1080
Qy	1023	TGTGGCCAACTGTGTTCACTGGAGCAAGAAAGAGATCTCATAGGACGAGGGGGAATGGT	1082
Db	1081	TGTGGCCAACTGTGTTCACTGGAGCAAGAAAGAGATCTCATAGGACGAGGGGGAATGGT	1140
Qy	1083	TTCCCTCCAAGCTTGGGTGAGTGTGTTAACTGCTTATCAGCTATTACAGCATCTCCATGG	1142
Db	1141	TTCCCTCCAAGCTTGGGTGAGTGTGTTAACTGCTTATCAGCTATTACAGCATCTCCATGG	1200
Qy	1143	TTTCTCATGAACCTCTGTGGTTTCATCTTCCTTCCTTAGTTGACCTGACAGCTTGGTT	1202
Db	1201	TTTCTCATGAACCTCTGTGGTTTCATCTTCCTTCCTTAGTTGACCTGACAGCTTGGTT	1260
Qy	1203	AGACCTAGATTTAACCCCTAAGTAAAGATGCTGGGGTATAGAACCTAAGAAATTTTCCCCC	1262
Db	1261	AGACCTAGATTTAACCCCTAAGTAAAGATGCTGGGGTATAGAACCTAAGAAATTTTCCCCC	1320
Qy	1263	AAGGACTCTTGCTTCTTAAAGCCCTTCTGGCTTCGTTTATGGCTTCATTAAAGATATAA	1322
Db	1321	AAGGACTCTTGCTTCTTAAAGCCCTTCTGGCTTCGTTTATGGCTTCATTAAAGATATAA	1380
Qy	1323	GCCTAACTTTGCTGCTAGTCTCAGGAGAAACCTTTAACCAACAAGTTTTTTATCATTTGAA	1382
Db	1381	GCCTAACTTTGCTGCTAGTCTCAGGAGAAACCTTTAACCAACAAGTTTTTTATCATTTGAA	1440
	1383	GACAATATTGAACAACCCCTATTCTTCTGGGATTTGAGAGGGGTGAATAGAGCCTTGAG	1442
bo	1441	GACAATATTGAACAACCCCTATTCTTCTGGGATTTGAGAGGGGTGAATAGAGCCTTGAG	1500
Qy	1443	ACTTTCCCTTTGTGGTAGGACTTTGGAGGAGAAATCCCTGGACTTTTCACTTAACCCCTCTG	1502
Db	1501	ACTTTCCCTTTGTGGTAGGACTTTGGAGGAGAAATCCCTGGACTTTTCACTTAACCCCTCTG	1560
Qy	1503	ACATACCTCCCAACACCCAGTTGATGGCTTTCCCGTAATAAAGATTCGGGATT	1554
Db	1561	ACATACCTCCCAACACCCAGTTGATGGCTTTCCCGTAATAAAGATTCGGGATT	1612

RESULT 7

US-09-902-853-206
; Sequence 206, Application US/09902853
; Publication No. US20020192659A1
-; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David

APPLICANT: Botstein, David

88.18; Score 1370; DB 9; Length 1620;

```

; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 206
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 973, 977, 996, 1003
; OTHER INFORMATION: unknown base
; US-09-902-853-206

```

Query Match

Best Local Similarity 92.7%; Pred. No. 0; Matches 1494; Conservative 0; Mismatches 4; Indels 114; Gaps 1;	
QY	57 AGATGGCGGCTTGGCACCTCTAATTTGCTCTCGTGATTCGGTGGCGGACACTTTCACGAT
DB	1 AGATGGCGGCTTGGCACCTCTAATTTGCTCTCGTGATTCGGTGGCGGACACTTTCACGAT
QY	117 GGTCGCCCAACCTTACTACTTCTGTGCGCCCTGTCTCTGTGTGCTTCCCTACTCGTGA
DB	61 GGTCGCCCAACCTTACTACTTCTGTGCGCCCTGTCTCTGTGTGCTTCCCTACTCGTGA
QY	177 GGAACCTGCCGCGCTCTGCCAGGCTCTGCCACCTCCCAACCGAAGAGCGGTAAACCCGTGTG
DB	121 GGAACCTGCCGCGCTCTGCCAGGCTCTGCCACCTCCCAACCGAAGAGCGGTAAACCCGTGTG
Y	237 ACTTTGACTGGAGAGAGTGGAGATCCTGATGTTTCTCAGTGCATTTGATGATGAAGA
JB	181 ACTTTGACTGGAGAGAGTGGAGATCCTGATGTTTCTCAGTGCATTTGATGATGAAGA
QY	297 ACCGCAGATCCA-----
DB	241 ACCGCAGATCCACTACTGTGGACAACATATAGCAACATTTTCATGTTTAGTAAAGTGG
QY	309 -----
DB	301 CCAACACAATTCCTTTCTTCGGCTTGGATATTGGCATGGCCCTACTTTACATCACACTCT
QY	309 -----TGGTCTCGATGACGTGCAAAACCCCTATATATGGCCCTGAGTATATCAAGT
DB	361 GCATAGTGTCTCGATGACGTGCNAACCCCTATATATGGCCCTGAGTATATCAAGT
QY	363 ACTTCAATGATAAAACCATTTGATGAGAACTAGAACGGGACAGAGGGTCACTTGGATTG
DB	421 ACTTCAATGATAAAACCATTTGATGAGAACTAGAACGGGACAGAGGGTCACTTGGATTG
QY	423 TGGAGTCTTTGGCAATTTGGTCTAAATGACTGCCAATCATTTGCCCTATCTATGCTGACC
DB	481 TGGAGTCTTTGGCAATTTGGTCTAAATGACTGCCAATCATTTGCCCTATCTATGCTGACC
QY	483 TCTCCCTTAAATACAACCTGTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGAGCGCTATA
DB	541 TCTCCCTTAAATACAACCTGTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGAGCGCTATA
Y	543 CTGATGTTAGTCGCGGTACAAAGTGAAGCATCAACCCCTACCAAGCAACTCCCTACCC
JB	601 CTGATGTTAGTCGCGGTACAAAGTGAAGCATCAACCCCTACCAAGCAACTCCCTACCC
QY	603 TGATCTGTTTCCAAAGTGGCAAGGCAATGGCGGGCCACAGATTGACAAGAAAGGAC
DB	661 TGATCTGTTTCCAAAGTGGCAAGGCAATGGCGGGCCACAGATTGACAAGAAAGGAC
QY	663 GGGCTGTCTCATGGACCTCTCTGAGGAGAAATGTGATCCGAGAAATTTAACTTAAATGAGC
DB	721 GGGCTGTCTCATGGACCTCTCTGAGGAGAAATGTGATCCGAGAAATTTAACTTAAATGAGC
QY	723 TATACCAAGCGGGCCAAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTG
DB	781 TATACCAAGCGGGCCAAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTG
QY	783 TGGCTTCAACCCCAACACAGTGTAGATGGGNAACCAAGAGGATTAATAAGATCCTC
DB	841 TGGCTTCAACCCCAACACAGTGTAGATGGGNAACCAAGAGGATTAATAAGATCCTC
QY	843 ACTTTGGCAGTGCTTCTCTCTCTCAATTTCCAGGCTCTTTCATATACCAACAGCCCTGAG
DB	901 ACTTTGGCAGTGCTTCTCTCTCTCAATTTCCAGGCTCTTTCATATACCAACAGCCCTGAG
QY	903 GCTGCAGCCCTTTATTTATCTTTTCCCTTTGGCTGTGACTGGGTGGGGCAGCATGCAGCT
DB	961 GCTGCAGCCCTTTNATTNATGTTTTCCTTTGGCTTGNACTGGNTGGGGCAGCATGCAGCT
QY	963 TCTGATTTTAAAGAGGCATCTAGGGAATTTGTACGACCCCTACAGGAAGCCCTGCCATGC

```

; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 206
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 973, 977, 996, 1003
; OTHER INFORMATION: unknown base
; US-09-907-824-206

Query Match      88.1%; Score 1370; DB 9; Length 1620;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1494; Conservative 0; Mismatches 4; Indels 114; Gaps 1;

; 57 AGATGGGGGCTTGGCACCTCTAATGCTCTCGTGATATTCGGTGCCTGCTTCCAGAT 116
; 1 AGATGGGGGCTTGGCACCTCTAATGCTCTCGTGATATTCGGTGCCTGCTTCCAGAT 60
; 117 GGCTCGCCCACTTACTACCTTCTGCGGCCCTGCTCTGCTGCTTCTCTACTCTGTA 176
; 61 GGCTCGCCCACTTACTACCTTCTGCGGCCCTGCTCTGCTGCTTCTCTACTCTGTA 120
; 177 GGAACCTGCCCGCTCTGCGCACGGTCTGCGCCACCCCAACGAGACGGTAACCGGTGTG 236
; 121 GGAACCTGCCCGCTCTGCGCACGGTCTGCGCCACCCCAACGAGACGGTAACCGGTGTG 180
; 237 ACTTTGACTGGAGAGAGTGGAGATCCGTGATGTTTCTCAGTGCCCATTTGTGATGAAGA 296
; 181 ACTTTGACTGGAGAGAGTGGAGATCCGTGATGTTTCTCAGTGCCCATTTGTGATGAAGA 240
; 297 ACCCCAGATCCA----- 308
; 241 ACCCGAGATCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAGTGG 300
; 309 ----- 308
; 301 CCAACACAATCTTTTCTCCCTTGGATATTCGCATGGGCTACTTTTACATCACACTCT 360
; 309 -----TGTCTCTGATGACGTGCAAAACCCCTTATATATGGCCCTGAGTATATCAAGT 362
; 361 GCATAGTGTCTCTGATGACGTGCAAAACCCCTTATATATGGCCCTGAGTATATCAAGT 420
; 363 ACTTCAANTGATAAAACCATTTGATGAGGAAGTGAACGGGACAGAGGGTCACTTGGATTG 422
; 421 ACTTCAANTGATAAAACCATTTGATGAGGAAGTGAACGGGACAGAGGGTCACTTGGATTG 480
; 423 TGGAGTCTCTTCCCAATTTGCTTAATGACTGCCAATCATTTGCCCTATCTATCTGCTGACC 482
; 481 TGGAGTCTCTTCCCAATTTGCTTAATGACTGCCAATCATTTGCCCTATCTATCTGCTGACC 540
; 483 TCTCCCTTAATATACAACTGTACAGGCTAAATTTTGGGAAGTGGATGTTGGACGCTATA 542
; 541 TCTCCCTTAATATACAACTGTACAGGCTAAATTTTGGGAAGTGGATGTTGGACGCTATA 600
; 543 CTGATGTTAGTACGGGTACAAAGTGAACACATCACCCCTCACCAAGCAACTCCCTTACC 602
; 601 CTGATGTTAGTACGGGTACAAAGTGAACACATCACCCCTCACCAAGCAACTCCCTTACC 660
; 603 TGATCCTGTTCACAGGTGGCAAGAGGCAATCGCGGCCACAGATTGACAAAGAGGAC 662
; 661 TGATCCTGTTCACAGGTGGCAAGAGGCAATCGCGGCCACAGATTGACAAAGAGGAC 720
; 663 GGGCTGTCTCATGAGACCTTCTCTGAGGAAATGTGATCCGAGAATTTAACTTTAAATGAGC 722
; 721 GGGCTGTCTCATGAGACCTTCTCTGAGGAAATGTGATCCGAGAATTTAACTTTAAATGAGC 780
; 723 TATACCAGGGGCCCAAGAACTATCAAGGCTGGAGACATATCCCTGAGGAGCAGCCTG 782
; 781 TATACCAGGGGCCCAAGAACTATCAAGGCTGGAGACATATCCCTGAGGAGCAGCCTG 840
; 783 TGGCTTCAACCCCAACACAGTGTGATGGGGAAGAAAGAGGATAAATAGATCCCTC 842
; 841 TGGCTTCAACCCCAACACAGTGTGATGGGGAAGAAAGAGGATAAATAGATCCCTC 900
; 843 ACTTTGGCAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 902
; 901 ACTTTGGCAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
; 903 GCTGACAGCTTTTATTTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 962
; 961 GCTGACAGCTTTTATTTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
; 963 TCTGATTTTAAAGAGGCATCTAGGGAATTTGTCAGGCACCTACAGGAAGCCCTGCCATGC 1022
; 1021 TCTGATTTTAAAGAGGCATCTAGGGAATTTGTCAGGCACCTACAGGAAGCCCTGCCATGC 1080
; 1023 TGTGCCCACACTGTTTCTACTGGAGCAAGAAAGATCTCTATAGGACGGGAGGGAATGGT 1082
; 1081 TGTGCCCACACTGTTTCTACTGGAGCAAGAAAGATCTCTATAGGACGGGAGGGAATGGT 1140
; 1083 TTCCCTCCCAAGCTTGGGTGAGTGTGTTTAACTGCTTATGAGCTTATTCAGACATCTCCATGG 1142
; 1141 TTCCCTCCCAAGCTTGGGTGAGTGTGTTTAACTGCTTATGAGCTTATTCAGACATCTCCATGG 1200
; 1143 TTTCTCCCATGAACACTCTGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1202
; 1201 TTTCTCCCATGAACACTCTGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
; 1203 AGACCTAGATTTAAACCTTAAGTAAGATGCTGGGTATAGAACGCTTAAGAAATTTTCCCCC 1262
; 1261 AGACCTAGATTTAAACCTTAAGTAAGATGCTGGGTATAGAACGCTTAAGAAATTTTCCCCC 1320
; 1263 AAGGACTCTCTCTTAAAGCCCTTCTGCTTCTGTTTATGGTCTTCTTATTAAGATATAA 1322
; 1321 AAGGACTCTCTCTTAAAGCCCTTCTGCTTCTGTTTATGGTCTTCTTATTAAGATATAA 1380
; 1323 GCCTAACTTGTGCTAGTCTCTAAGGAGAAACCTTTAACCACAAAGTTTATTCATCAATGAA 1382
; 1325 ----- 1380
```

1381	GCCTAACTTTGTCGCTAGTCCCTAAGGAGAAACCTTTAAACCAAAAGTTTTTATCATTTGAA	1440
Db		
1383	GACAATATTGAACAACACCCCTATTATTTGTGGGGATTGAGAAAGGGGTGAATAGAGGCTTGAG	1442
QY		
1441	GACAATATTGAACAACACCCCTATTATTTGTGGGGATTGAGAGGGGTGNAATAGAGGCTTGAG	1500
Db		
1443	ACTTTCCTTTGTGTGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCCTACTAACCCCTCTG	1502
QY		
1501	ACTTTCCTTTGTGTGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCCTACTAACCCCTCTG	1560
Db		
1503	ACATACTCCCCACACCCAGTTGTATGGCTTTCCCGTAATAAAAGATTGGGATT	1554
QY		
1561	ACATACTCCCCACACCCAGTTGTATGGCTTTCCCGTAATAAAAGATTGGGATT	1612
Db		

9. T.S.I.

```

001 3
1 Sequence 206, Application US/09907841
2 Publication No. US20020198366A1
3 GENERAL INFORMATION:
4 APPLICANT: Genentech, Inc.
5 APPLICANT: Ashkenazi, Avi
6 APPLICANT: Botstein, David
7 APPLICANT: Desnoyers, Luc
8 APPLICANT: Eaton, Dan L.
9 APPLICANT: Ferrara, Napoleone
10 APPLICANT: Filvaroff, Ellen
11 APPLICANT: Fong, Sherman
12 APPLICANT: Gao, Wei-Qiang
13 APPLICANT: Gerber, Hanspeter
14 APPLICANT: Gerritsen, Mary E.
15 APPLICANT: Goddard, A.
16 APPLICANT: Godowski, Paul J.
17 APPLICANT: Grimaldi, Christopher J.
18 APPLICANT: Gurney, Austin L.
19 APPLICANT: Hillan, Kenneth, J.
20 APPLICANT: Kljavin, Ivar J.
21 APPLICANT: Mather, Jennie P.
22 APPLICANT: Pan, James
23 APPLICANT: Paoni, Nicholas F.
24 APPLICANT: Roy, Margaret Ann
25 APPLICANT: Stewart, Timothy A.
26 APPLICANT: Tumas, Daniel
27 APPLICANT: Williams, P. Mickey
28 APPLICANT: Wood, William, I.
29 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
30 FILE OF INVENTION: Acids Encoding the Same
31 FILE REFERENCE: 10466-14
32 CURRENT APPLICATION NUMBER: US/09/907,841
33 PRIOR FILING DATE: 2001-11-20
34 PRIOR APPLICATION NUMBER: PCT/US00/04414
35 PRIOR FILING DATE: 2000-02-22
36 PRIOR APPLICATION NUMBER: US 60/143,048
37 PRIOR FILING DATE: 1999-07-07
38 PRIOR APPLICATION NUMBER: US 60/145,698
39 PRIOR FILING DATE: 1999-07-26
40 PRIOR APPLICATION NUMBER: US 60/146,222
41 PRIOR FILING DATE: 1999-07-28
42 PRIOR APPLICATION NUMBER: PCT/US99/20594
43 PRIOR FILING DATE: 1999-09-08
44 PRIOR APPLICATION NUMBER: PCT/US99/20944
45 PRIOR FILING DATE: 1999-09-13
46 PRIOR APPLICATION NUMBER: PCT/US99/21090
47 PRIOR FILING DATE: 1999-09-15
48 PRIOR APPLICATION NUMBER: PCT/US99/21547
49 PRIOR FILING DATE: 1999-09-15
50 PRIOR APPLICATION NUMBER: PCT/US99/23089
51 PRIOR FILING DATE: 1999-10-05
52 PRIOR APPLICATION NUMBER: PCT/US99/28214
53 PRIOR FILING DATE: 1999-11-29
54 Remaining Prior Application data removed - See File Wrapper or PALM.
55 NUMBER OF SEQ ID NOS: 423
56 SEQ ID NO 206

```

NUMBER OF SEASONS
SEO ID NO 206

	: LENGTH: 1620	
	: TYPE: DNA	
	: ORGANISM: Homo sapiens	
	: FEATURE:	
	: NAME/KEY: modified_base	
	: LOCATION: (973)	
	: OTHER INFORMATION: a, t, c or g	
	: FEATURE:	
	: NAME/KEY: modified_base	
	: LOCATION: (977)	
	: OTHER INFORMATION: a, t, c or g	
	: FEATURE:	
	: NAME/KEY: modified_base	
	: LOCATION: (996)	
	: OTHER INFORMATION: a, t, c or g	
	: FEATURE:	
	: NAME/KEY: modified_base	
	: LOCATION: (1003)	
	: OTHER INFORMATION: a, t, c or g	
	US-09-907-841-206	
	Query Match 88.1%; Score 1370; DB 9; Length 1620;	
	Best Local Similarity 92.7%; Pred. No. 0;	
	Matches 1494; Conservative 0; Mismatches 4; Indels 114; Gaps 1;	
Qy	57 AGATGGCGGCTTTGGACACTCTAATTGCCTCTCGTGATTCGGTGGCGGCACTTTCAGCAT 116	
Db		
	1 AGATGGCGGCTTTGGCACCTCTAATTGCTCTCGTGATTCGGTGGCGGCACTTTCAGCAT 60	
Qy	117 GGCTGCGCCAACTTACTACTCTGTGCGGCCCTGCTCTCTGCTGCTTCCCTACTCTCGTGA 176	
Db		
	61 GGCTGCGCCAACTTACTACTCTGTGCGGCCCTGCTCTCTGCTGCTTCCCTACTCTCGTGA 120	
Qy	177 GGAACATGCGCGCGCTCTGCCACGGTGTCCCCACCACAACGCAAGCGTAACCCGTGTG 236	
Db		
	121 GGAACATGCGCGCGCTCTGCCACGGTGTGCCACCCAACGCAAGCGTAACCCGTGTG 180	
Qy	237 ACTTTGACTGGAGAGAAGTGGAGATCCTGATGTTTTCTCAGTGCCATTTGTGATGATGAAGA 296	
Db		
	181 ACTTTGACTGGAGAGAAGTGGAGATCCTGATGTTTTCTCAGTGCCATTTGTGATGATGAAGA 240	
Qy	297 ACGCAGATGCA ----- 308	
Db		
	241 ACCGCAGATCCATCACTGTCGGACACATATAGGCAACATTTTCATGTTTTAGTAAAGTGG 300	
Qy	309 ----- 308	
Db	301 CCAACACAATTTTCTTCGGCTTGGATATTCGCATGGCCCTACTTTACATCACACTCT 360	
Qy	309 -----TGTTCCTGATGACGTGAAAACCCCCCCTATATATATGGCCCTGAGTATATCAAAGT 362	
Db		
	361 GCATAGTGTTCCTGATGACGTGAAAACCCCCCCTATATATGGCCCTGAGTATATCAAAGT 420	
Qy	363 ACTTCAATGATAAAACCATTTGATGAGGAAC TAGAACGGGCAAGAGGGTCACATTGGATTG 422	
Db		
	421 ACTTCAATGATAAAACCATTTGATGAGGAAC TAGAACGGGCAAGAGGGTCACATTGGATTG 480	
Qy	423 TGGAGTCTTTGGCAATTTGGTCTAATGACTGCGCAATCATTTGGCCCTATCTATGCTGACCC 482	
Db		
	481 TGGAGTCTTTGGCAATTTGGTCTAATGACTGCGCAATCATTTGGCCCTATCTATGCTGACCC 540	
Qy	483 TCCTCCCTTAATAACAACGTGACAGGGCTAAATTTTGGGAAGGTGGATGTTGGGACGCTATA 542	
Db		
	541 TCCTCCCTTAATAACAACGTGACAGGGCTAAATTTTGGGAAGGTGGATGTTGGGACGCTATA 600	
Qy	543 CTGATGTTAGTAGCGGTACAAAAGTAGGCACATCACTCCCTCACCAAGCAACTCCCTACCC 602	
Db		
	601 CTGATGTTAGTAGCGGTACAAAAGTAGGCACATCACTCCCTCACCAAGCAACTCCCTACCC 660	
Qy	603 TGATCTCTGTTCCAAAGTGGCAAGAGGCAATCGCGGGGCCACAGATTGACAGAAAGGAC 662	
Db		
	661 TGATCTCTGTTCCAAAGTGGCAAGAGGCAATCGCGGGGCCACAGATTGACAGAAAGGAC 720	

QY 663 GGCTCTCTCATGGACCTTCTCTCAGGAGAAATGTGATCCGAGATTTAACTTAATGAGC 722
Db 721 GGCTCTCTCATGGACCTTCTCTCAGGAGAAATGTGATCCGAGATTTAACTTAATGAGC 780
QY 723 TATACAGCGGGCCAAAGAACTATCAAGAGCTGGAGACAATATCCCTGAGGAGCAGCTG 782
Db 781 TATACAGCGGGCCAAAGAACTATCAAGAGCTGGAGACAATATCCCTGAGGAGCAGCTG 840
QY 783 TGGCTTCAACCCCCACACAGTGTGAGATGGGAAAAACAAGAGGATAAATAAGATCCTC 842
Db 841 TGGCTTCAACCCCCACACAGTGTGAGATGGGAAAAACAAGAGGATAAATAAGATCCTC 900
QY 843 ACTTTGGCAGTGTCTCTCTCTCTCAATTCAGGCTTTCATACCAACCAACGCTGAG 902
Db 901 ACTTTGGCAGTGTCTCTCTCTCTCAATTCAGGCTTTCATACCAACCAACGCTGAG 960
QY 903 GCTGACGCTTTTATTTATTTTCCCTTTGGCTGTGACTGGGTGGGCGCAGCATGAGCT 962
Db 961 GCTGACGCTTTTATTTATTTTCCCTTTGGCTGTGACTGGGTGGGCGCAGCATGAGCT 1020
QY 963 TGTGATTTTAAAGAGCATCTAGGAAATGTGAGGACCCCTACAGGAAGGCTGCCATGC 1022
Db 1021 TGTGATTTTAAAGAGCATCTAGGAAATGTGAGGACCCCTACAGGAAGGCTGCCATGC 1080
1023 TGTGGCCAACTGTTTCACTGAGCAACAAAGAGATCTCATAGGACGAGGGGAAATGGT 1082
Db 1081 TGTGGCCAACTGTTTCACTGAGCAACAAAGAGATCTCATAGGACGAGGGGAAATGGT 1140
QY 1083 TTCCCTCCAAAGCTTGGTCACTGTGTTAACTGCTTATCAGCTATTCAGACATCTCCATGG 1142
Db 1141 TTCCCTCCAAAGCTTGGTCACTGTGTTAACTGCTTATCAGCTATTCAGACATCTCCATGG 1200
QY 1143 TTTCTCCATGAACACTCTGTGGTTTCATCTCTTCTTGTGACCTGACAGCTTGGTT 1202
Db 1201 TTTCTCCATGAACACTCTGTGGTTTCATCTCTTCTTGTGACCTGACAGCTTGGTT 1260
QY 1203 AGACCTAGATTTTAACTTAAAGTAAGATGCTGGGTATAGAACCTTAAGATTTTCCCTC 1262
Db 1261 AGACCTAGATTTTAACTTAAAGTAAGATGCTGGGTATAGAACCTTAAGATTTTCCCTC 1320
QY 1263 AGGACTCTTGTCTTAAAGCCCTTCTGCTTATGCTTCTTATGCTTCTTAAAGATATAA 1322
Db 1321 AGGACTCTTGTCTTAAAGCCCTTCTGCTTATGCTTCTTATGCTTCTTAAAGATATAA 1380
QY 1323 GCTTAACCTTGTGCTAGTCTTAAAGGAGAACCTTTAACCAACAAGTTTATCATGTA 1382
Db 1381 GCTTAACCTTGTGCTAGTCTTAAAGGAGAACCTTTAACCAACAAGTTTATCATGTA 1440
QY 1383 GACAATATTGAACACCCCTATTTTGTGGGATTTGAGAGGGGTGAATAGAGCTTGTAG 1442
Db 1441 GACAATATTGAACACCCCTATTTTGTGGGATTTGAGAGGGGTGAATAGAGCTTGTAG 1500
1443 ACTTTCTTGTGTGTGAGTCTTGGAGAGAAATCCCTGGACTTTTCACTAACCCCTCTG 1502
Db 1501 ACTTTCTTGTGTGTGAGTCTTGGAGAGAAATCCCTGGACTTTTCACTAACCCCTCTG 1560
QY 1503 ACATCTCCCAACCCAGTGTGCTTCCGTAATAAAGATTTGGGATTT 1554
Db 1561 ACATCTCCCAACCCAGTGTGCTTCCGTAATAAAGATTTGGGATTT 1612

RESULT 10

US-09-904-011-206
; Sequence 206, Application US/09904011
; Publication No. US20030003530A1

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 206
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 973, 977, 996, 1003
; OTHER INFORMATION: unknown base
US-09-904-011-206

Query Match

Best Local Similarity 88.1%; Score 1370; DB 9; Length 1620;
Matches 1494; Conservativity 92.7%; Pred.No. 0;
Indels 114; Gaps 1;

QY	57	AGATGGCGGTC	TTGGCACCTCTAAT	TGCTCTCTGTTATTCGGTCCCGACT	TTTCACGAT	116
DB	1	AGATGGCGGTC	TTGGCACCTCTAAT	TGCTCTCTGTTATTCGGTCCCGACT	TTTCACGAT	60
QY	117	GGCTCGCCCAAC	CTTACTACTCTCTG	CGCCCTGCTCTCTGCTGCGCTT	CCTACTCGTGA	176
DB	61	GGCTCGCCCAAC	CTTACTACTCTCTG	CGCCCTGCTCTCTGCTGCGCTT	CCTACTCGTGA	120
QY	177	GGAACCTGCCCG	CGCTCTGCGACGGT	CTGCCACCCCAACGCGGTAACCGGTG	236	
DB	121	GGAACCTGCCCG	CGCTCTGCGACGGT	CTGCCACCCCAACGCGGTAACCGGTG	180	
QY	237	ACTTTGACTGG	GAGAAAGTGGAGAT	CTGATGTTTCTCAGTGCCATTG	TGATGATGAAGA	296
DB	181	ACTTTGACTGG	GAGAAAGTGGAGAT	CTGATGTTTCTCAGTGCCATTG	TGATGATGAAGA	240
QY	297	ACCGCAGATCCA	-----	-----	-----	308
DB	241	ACCGCAGATCCA	CTCTGTGGAGCAACAT	TATAGCAACATTTTCATGTT	TAGTAAAGTGG	300
QY	309	-----	-----	-----	-----	308
DB	301	CCAACAAATCT	TTTCTTCGGCTTG	GATATTCGCATGGGCTACT	TTTACATCACACTCT	360
QY	309	-----	-----	-----	-----	362
DB	361	GCATAGTGTTC	CTGATGACGTGCA	AAACCCGCCCTATATATGGGCCCT	TGAGTATCAAGT	420
QY	363	ACTTCAATGATA	AAACCAATTCATG	CAGAACTAGAACGGGACACAGAGGGT	CACTTGGATTG	422
DB	421	ACTTCAATGATA	AAACCAATTCATG	CAGAACTAGAACGGGACACAGAGGGT	CACTTGGATTG	480
QY	423	TGGAGTCTTTGG	CAATTTGGTCTAAT	GACTGCCAATCATTTGCCCTATCT	TATGCTGACC	482
DB	481	TGGAGTCTTTGG	CAATTTGGTCTAAT	GACTGCCAATCATTTGCCCTATCT	TATGCTGACC	540
QY	483	TCCTCCCTTAAT	ACAACTGTACAGG	GCTAAATTTTGGGAAGGTGGATGTT	TGGACGCTATA	542
DB	541	TCCTCCCTTAAT	ACAACTGTACAGG	GCTAAATTTTGGGAAGGTGGATGTT	TGGACGCTATA	600
QY	543	CTGATGTTAGT	ACGGGTACAAAGT	GAGCACATCACCCCTCACCAAGCACT	CCCTACCC	602
DB	601	CTGATGTTAGT	ACGGGTACAAAGT	GAGCACATCACCCCTCACCAAGCACT	CCCTACCC	660
QY	603	TGATCTCTGTTCC	AAAGTGGCAAGAGG	CGCATGCGCGGCCACAGATTG	ACAAGAAAGAC	662
DB	661	TGATCTCTGTTCC	AAAGTGGCAAGAGG	CGCATGCGCGGCCACAGATTG	ACAAGAAAGAC	720
QY	663	GGCGTGCTCAT	GGACCTTCTCTGAG	GAGATGTGATCCGAGATTTAACT	TAAATGAGC	722
DB	721	GGCGTGCTCAT	GGACCTTCTCTGAG	GAGATGTGATCCGAGATTTAACT	TAAATGAGC	780
QY	723	TATACACGCGGG	CCAGAAACTATCA	AAGCTGGAGACAATATCCCT	TGAGGACGAGCCTG	782
DB	781	TATACACGCGGG	CCAGAAACTATCA	AAGCTGGAGACAATATCCCT	TGAGGACGAGCCTG	840
QY	783	TGGCTTCAACCC	CCACCAACAGTGT	CAGATGGGGAAAAACAAGAGGATA	ATAAGATCCTC	842
DB	841	TGGCTTCAACCC	CCACCAACAGTGT	CAGATGGGGAAAAACAAGAGGATA	ATAAGATCCTC	900
QY	843	ACTTTGGCAGT	GTCTCTCTCTCAAT	TCCAGGCTTTTCCATATACCAAC	CAAGCCTGAG	902
DB	901	ACTTTGGCAGT	GTCTCTCTCTCAAT	TCCAGGCTTTTCCATATACCAAC	CAAGCCTGAG	960
QY	903	GCTCCAGCCTTT	TATTTATGTTTCC	CTTTGGTGTGACTGGGTGGGCGC	AGCATGCAAGT	962
DB	961	GCTCCAGCCTTT	TATTTATGTTTCC	CTTTGGTGTGACTGGGTGGGCGC	AGCATGCAAGT	1020
QY	963	CTTGATTTTAA	GAGGCATCTAGG	GAATGTTCAGGCAACCTTACAGGAAG	CGCCTGCCATGC	1022
DB	1021	CTTGATTTTAA	GAGGCATCTAGG	GAATGTTCAGGCAACCTTACAGGAAG	CGCCTGCCATGC	1080

RESULT 11

US-10-174-590-31
; Sequence 31, Application US/10174590

; PUBLICATION NO: 0320
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3430R1C42

; CURRENT APPLICATION NUMBER: US/10/174,590

; CURRENT FILING DATE: 2002-06-18

; prior application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 31

```

; SEQ ID NO 31
: LENGTH: 1630

```

```

; LENGTH: 1620
; TYPE: DNA

```

TYPE: DNA
ORGANISM: Homo sapiens

```

; ORGANISM: Homo sapien
. FEATURE.

```

```
, FEATURE:  
: NAME / KEY. insure
```

; NAME/KEY: UNISURE
: LOCATION: 973. 97

LOCATION: 573, 577, 558, 1003
OTHER INFORMATION: unknown ba

US-10-174-590-31

1
2
3
4
5
6
7
8
9
10
11
12

Query Match

Best Local Similarity 92.78; pred. No. 0;

Best local similarity	52.7%	1160	NO: 0
Matches 1494;	Conservative	0:	Mismatches
		4:	Indels
			114: Gaps
			1:

[illegible]

QY 57 AGATGGCGGTCTTGGCACCTCTAATTGCTCTCTGTGTATTCGGTGGCGGACCTTTCAGGAT 116
Db 1 AGATGGCGGTCTTGGCACCTCTAATTGCTCTCTGTGTATTCGGTGGCGGACCTTTCAGGAT 60
QY 117 GGTCTGCCCAACCTTACTTCTCTGCTGCGCCCTGCTCTCTGCTGCTTCTTCTACTCTGTGA 176
Db 61 GGTCTGCCCAACCTTACTTCTCTGCTGCGCCCTGCTCTCTGCTGCTTCTTCTACTCTGTGA 120
QY 177 GGAACCTGCGCGGTCTTCTGCCAGGTCTGCCACCCCAACGCAAGCGGTAACCCCGTGTG 236
Db 121 GGAACCTGCGCGGTCTTCTGCCAGGTCTGCCACCCCAACGCAAGCGGTAACCCCGTGTG 180
QY 237 ACTTTGACTGGAGAGAGTGGAGATCCTGTATGTTTCTCAGTGCATGTGATGATGAGA 296
Db 181 ACTTTGACTGGAGAGAGTGGAGATCCTGTATGTTTCTCAGTGCATGTGATGATGAGA 240
QY 297 ACCGAGATGCCA----- 308
Db 241 ACCGAGATGCCA----- 300
QY 309 ----- 308
Db 301 CCAACACAATCTTTTCTTCCGCTTGGATATTCGCATGGCCCTTACATCACACTCT 360
QY 309 -----TGTCTCTGATGAGTGCACACCCCTCTATATATGCGCCCTGAGTATATCAAGT 362
Db 361 GCATAGTGTCTCTGATGAGTGCACACCCCTCTATATATGCGCCCTGAGTATATCAAGT 420
QY 363 ACTTCAATGATAAACCATTGATGAGGAAGTACAGCGGCAAGAGGGTCACTTGGATTG 422
Db 421 ACTTCAATGATAAACCATTGATGAGGAAGTACAGCGGCAAGAGGGTCACTTGGATTG 480
QY 423 TGGAGTCTTTGCAATTTGGTCTTAATGACTGCGCAATCTTTGCCCTCTATCTATGCTGACC 482
Db 481 TGGAGTCTTTGCAATTTGGTCTTAATGACTGCGCAATCTTTGCCCTCTATCTATGCTGACC 540
QY 483 TCTCCCTTAATACACTGTACAGGGCTAAATTTGGGAAGGTGGATTTGGAGCGCTATA 542
Db 541 TCTCCCTTAATACACTGTACAGGGCTAAATTTGGGAAGGTGGATTTGGAGCGCTATA 600
QY 543 CTGATCTTAGTACGGGTACAAAGTGCAGCATACCCCTCACCACAGCACTCCCTACCC 602
Db 601 CTGATCTTAGTACGGGTACAAAGTGCAGCATACCCCTCACCACAGCACTCCCTACCC 660
QY 603 TGATCCTGTTCCTCAAGTGGCAAGGAGCAATGCGGGCCACAGATTTGACAAAGAGGAC 662
Db 661 TGATCCTGTTCCTCAAGTGGCAAGGAGCAATGCGGGCCACAGATTTGACAAAGAGGAC 720
QY 663 GGGCTGTCTCATGGACCTTCTCTGAGGAGAAATGTATCGGAGAAATTAATTAATGAGC 722
Db 721 GGGCTGTCTCATGGACCTTCTCTGAGGAGAAATGTATCGGAGAAATTAATTAATGAGC 780
QY 723 TATACAGCGGGCCAAAGAACTATCAAGGCTGGAGCAATATCCCTGAGGAGCAGCCTG 782
Db 781 TATACAGCGGGCCAAAGAACTATCAAGGCTGGAGCAATATCCCTGAGGAGCAGCCTG 840
QY 783 TGGCTTCAACCCCCACAGAGTGTACAGTGGGAAACAAAGAGGATAAATAAGATCCTC 842
Db 841 TGGCTTCAACCCCCACAGAGTGTACAGTGGGAAACAAAGAGGATAAATAAGATCCTC 900
QY 843 ACTTTGGCAGTCTTCTCTCTCTCAATTCAGGCTCTTTCCATACACACAGCCTGAG 902
Db 901 ACTTTGGCAGTCTTCTCTCTCTCAATTCAGGCTCTTTCCATACACACAGCCTGAG 960
QY 903 GCTGAGCCTTTTATTTATTTTCTTCTTGGCTGTGACTGGGTGGGCGGAGCATGAGCT 962
Db 961 GCTGAGCCTTTTATTTATTTTCTTCTTGGCTGTGACTGGGTGGGCGGAGCATGAGCT 1020
QY 963 TCTGATTTTAAAGAGCATCTAGGGAATTTGTCAGGCAACCCCTACAGGAAGCCTTGCATGC 1022
Db 1021 TCTGATTTTAAAGAGCATCTAGGGAATTTGTCAGGCAACCCCTACAGGAAGCCTTGCATGC 1080

QY 1023 TGTGGCCAACTGTTTCACTGGAGCAAGAGAGATCTCATAGGACGAGGGGAAATGGT 1082
Db 1081 TGTGGCCAACTGTTTCACTGGAGCAAGAGAGATCTCATAGGACGAGGGGAAATGGT 1140
QY 1083 TTTCCCTCCAAAGCTTGGGTCACTGTTAACTGCTTATCAGCTATTCAGACATCTCCATGG 1142
Db 1141 TTTCCCTCCAAAGCTTGGGTCACTGTTAACTGCTTATCAGCTATTCAGACATCTCCATGG 1200
QY 1143 TTTCTCCATGAACACTCTGTGGTTTTCATCTTCTTATGACCTGACAGCTTGGTT 1202
Db 1201 TTTCTCCATGAACACTCTGTGGTTTTCATCTTCTTATGACCTGACAGCTTGGTT 1260
QY 1203 AGACCTAGATTTAAACCTTAAGGTAAGATGCTGGGTATAGAAGCTAAGAAATTTTCCCCC 1262
Db 1261 AGACCTAGATTTAAACCTTAAGGTAAGATGCTGGGTATAGAAGCTAAGAAATTTTCCCCC 1320
QY 1263 AAGGACTCTTGTCTTAAAGCCCTTCTGCTTCTGTTTATGGTCTTCAATTAAGATATA 1322
Db 1321 AAGGACTCTTGTCTTAAAGCCCTTCTGCTTCTGTTTATGGTCTTCAATTAAGATATA 1380
QY 1323 GCCTAATTTGTGCTAGTCTTAAAGGAGAAACCTTTAAACCAAAAGTTTATCATATTGAA 1382
Db 1381 GCCTAATTTGTGCTAGTCTTAAAGGAGAAACCTTTAAACCAAAAGTTTATCATATTGAA 1440
QY 1383 GACAATATTGAACAAACCCCTATTTTGTGGGATTTGAGAGGGGTGAATAGAGCTTGAG 1442
Db 1441 GACAATATTGAACAAACCCCTATTTTGTGGGATTTGAGAGGGGTGAATAGAGCTTGAG 1500
QY 1443 ACTTCTCTTGTGTGGTAGGACTTGGAGGAGAAATCCCTCGGACTTTTCACTAACCCCTCTG 1502
Db 1501 ACTTCTCTTGTGTGGTAGGACTTGGAGGAGAAATCCCTCGGACTTTTCACTAACCCCTCTG 1560
QY 1503 ACATCTCCCAACCCAGTTGATGCTTTCCGTAATAAAGATTTGGGATT 1554
Db 1561 ACATCTCCCAACCCAGTTGATGCTTTCCGTAATAAAGATTTGGGATT 1612

RESULT 14

US-09-906-742-206
; Sequence 206, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,742
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350

```

; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 206
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 973, 977, 996, 1003
; OTHER INFORMATION: unknown base
; 09-906-742-206

Query Match      88.1%; Score 1370; DB 9; Length 1620;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1494; Conservative 0; Mismatches 4; Indels 114; Gaps 1;

Oy 57 AGATGGCGGTCTTGGACCTCTTAATTGCTCTCGTGATTCGGTGGCGGACCTTCACGAT 116
Db 1 AGATGGCGGTCTTGGACCTCTTAATTGCTCTCGTGATTCGGTGGCGGACCTTCACGAT 60

Oy 117 GGCTCGCCCAACTTACTACCTCTCTCGGCCCTCTCTCTGCTGCTTCCCTACTCGTGA 176
Db 61 GGCTCGCCCAACTTACTACCTCTCTCGGCCCTCTCTCTGCTGCTTCCCTACTCGTGA 120

Oy 177 GAAACTGCGCGGCTCTGCCAGGTCTGCCACCCCAACGCGGAAGACGGTAACCCGGTG 236
Db 121 GAAACTGCGCGGCTCTGCCAGGTCTGCCACCCCAACGCGGAAGACGGTAACCCGGTG 180

Oy 237 ACTTFFACTGGAGAGAGTGGAGATCCCTGATCTTCTCAGTGCCCATTTGATGATGAAGA 296
Db 181 ACTTFFACTGGAGAGAGTGGAGATCCCTGATCTTCTCAGTGCCCATTTGATGATGAAGA 240

Oy 297 ACCGAGATCCA----- 308
Db 241 ACCGAGATCCA----- 300

Oy 309 ----- 308
Db 301 CCAACACAATCTTTTCTCCGCTTGGATATTTCGATGGCGCTACTTTTACATCACACTCT 360

309 QY -----TGTTCCTGATGACGTGCAAAACCCCTATATATATGGCCCTGAGTATATCAAGT 362
361 Db GCATAGTGTTCCTGATGACGTGCAAAACCCCTATATATATGGCCCTGAGTATATCAAGT 420
363 QY ACTTCAATGATAAAACCATTTGATGAGAACTAGAACGGGACAAAGAGGGTCACTTGGATTG 422
421 Db ACTTCAATGATAAAACCATTTGATGAGAACTAGAACGGGACAAAGAGGGTCACTTGGATTG 480
423 QY TGGAGTTCTTTGGCAATTTGGTCTAATGACTGCCAATCAATTTGGCCCTATCTATGCTGACC 482
481 Db TGGAGTTCTTTGGCAATTTGGTCTAATGACTGCCAATCAATTTGGCCCTATCTATGCTGACC 540
483 QY TCTCCCTTAAATACAACCTGTACAGGGCTAAATTTTGGAAAGGTGGATTTGGAGGCTATA 542
541 Db TCTCCCTTAAATACAACCTGTACAGGGCTAAATTTTGGAAAGGTGGATTTGGAGGCTATA 600
543 QY CTGATGTTAGTAGCGGTACAAAGTAGGACACATCAACCCCTCACCAAGCAACTCCCTACCC 602
601 Db CTGATGTTAGTAGCGGTACAAAGTAGGACACATCAACCCCTCACCAAGCAACTCCCTACCC 660
603 QY TGATCCTGTTCCAAAGTGGCAAGGAGGCAATGCGGGGCCACAGATTGACAAAGAGGAC 662
661 Db TGATCCTGTTCCAAAGTGGCAAGGAGGCAATGCGGGGCCACAGATTGACAAAGAGGAC 720
663 QY GGGCTGCTCATGGACCTTCTCTGAGGAGAAATGATCCGAGAAATTAATTAATGAGC 722
721 Db GGGCTGCTCATGGACCTTCTCTGAGGAGAAATGATCCGAGAAATTAATTAATGAGC 780
723 QY TATACCAGCGGGCCAAAGAACTATCAAAAGCTGGAGACAATATCCCTGAGGACGACGCTG 782
781 Db TATACCAGCGGGCCAAAGAACTATCAAAAGCTGGAGACAATATCCCTGAGGACGACGCTG 840
783 QY TGCTTTCAACCCCCACCACAGTGTGATGGGAAACCAAGAGAGATAAATAAGATCCTC 842
841 Db TGCTTTCAACCCCCACCACAGTGTGATGGGAAACCAAGAGAGATAAATAAGATCCTC 900
843 QY ACTTTGGCAGTGTCTCTCTCTCTCAATTCAGGCTCTTTCCATACCACAGCCTGAG 902
901 Db ACTTTGGCAGTGTCTCTCTCTCTCAATTCAGGCTCTTTCCATACCACAGCCTGAG 960
903 QY GCTGACGCTTTTATTTATTTTCCCTTTGGCTGTGACTGGTGGGGGACGACGTGACGT 962
961 Db GCTGACGCTTTTATTTATTTTCCCTTTGGCTGTGACTGGTGGGGGACGACGTGACGT 1020
963 QY TCTGATTTTAAAGAGGACATCTAGGGAATGTCAGGCACCCCTACAGGAAGCCCTGCCATGC 1022
1021 Db TCTGATTTTAAAGAGGACATCTAGGGAATGTCAGGCACCCCTACAGGAAGCCCTGCCATGC 1080
1023 QY TGTGGCCAACTGTTTCACTGGAGCAAGAAAGATCTCATAGGACGGAGGGGAAATGGT 1082
1081 Db TGTGGCCAACTGTTTCACTGGAGCAAGAAAGATCTCATAGGACGGAGGGGAAATGGT 1140
1083 QY TTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATCAGTATTTCAGACATCTCCATGG 1142
1141 Db TTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATCAGTATTTCAGACATCTCCATGG 1200
1143 QY TTTCTCCATGAAACTCTGTGGTTTTCATCTTCTTTAGTTGACCTGCACACGCTTGGTT 1202
1201 Db TTTCTCCATGAAACTCTGTGGTTTTCATCTTCTTTAGTTGACCTGCACACGCTTGGTT 1260
1203 QY AGACCTAGATTTTAAACCTTAAGGTAAGATGCTGGGGTATAGACGCTAAGAAATTTTCCCC 1262
1261 Db AGACCTAGATTTTAAACCTTAAGGTAAGATGCTGGGGTATAGACGCTAAGAAATTTTCCCC 1320
1263 QY AGGAGCTCTTGTCTTAAAGCCCTTCTGCTTCTGCTTCTGCTTCTTCTTCTTCTTCTTCTT 1322
1321 Db AGGAGCTCTTGTCTTAAAGCCCTTCTGCTTCTGCTTCTGCTTCTTCTTCTTCTTCTTCTT 1380
1323 QY GCTTAACCTTTGTGCTAGTCCTTAAGGAGAAACCTTTAAACCAACAAAGCTTTTATCATTTGAA 1382
1381 Db GCTTAACCTTTGTGCTAGTCCTTAAGGAGAAACCTTTAAACCAACAAAGCTTTTATCATTTGAA 1440
```

```
Qy 1383 GACAATATTGAACACCCCTATTGTTGGGATTGAGAAGGGTGAATAGAGGCTTGAG 1442
Db 1441 GACANTATTGAACACCCCTATTGTTGGGATTGAGAAGGGTGAATAGAGGCTTGAG 1500
Qy 1443 ACTTTCCTTTGTGTGGTAGGACTTTGGAGGAGAAATCCCTGGACTTTTCACTAAACCCCTCTG 1502
Db 1501 ACTTTCCTTTGTGTGGTAGGACTTTGGAGGAGAAATCCCTGGACTTTTCACTAAACCCCTCTG 1560
Qy 1503 ACATACCTCCCAACACCCAGTTGATGGCTTTCCTGTAATAAAGATTTGGGATT 1554
Db 1561 ACATACCTCCCAACACCCAGTTGATGGCTTTCCTGTAATAAAGATTTGGGATT 1612

RESULT 15
US-10-173-706-31
; Sequence 31, Application US/10173706
; Publication No. US2003002293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 31
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 973..977, 996, 1003
; OTHER INFORMATION: unknown base

US-10-173-706-31
Query Match 88.1%; Score 1370; DB 9; Length 1620;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1494; Conservative 0; Mismatches 4; Indels 114; Gaps 1;

Qy 57 AGATGGCGGTCTGGCACCTCTAATTGCTCTCGTGTATTCGGTGGCGGACTTTTCACGAT 116
1 AGATGGCGGTCTGGCACCTCTAATTGCTCTCGTGTATTCGGTGGCGGACTTTTCACGAT 60
Qy 117 GGCTCGCCCAACCTTACTACCTCTGTCGGCCCTGCTCTGCTGCTTCTACTCCTGTA 176
61 GGCTCGCCCAACCTTACTACCTCTGTCGGCCCTGCTCTGCTGCTTCTACTCCTGTA 120
Qy 177 GGAACCTGCCCGCTCTGCCACGGTCTGCCACCCCAACCGGAGACGGTAACCGGTGTG 236
121 GGAACCTGCCCGCTCTGCCACGGTCTGCCACCCCAACCGGAGACGGTAACCGGTGTG 180
Qy 237 ACTTTGACTGGAGAGAGTGAGATCCCTGATGTTTCTCAGTGCCCATTTGATGATGAAGA 296
181 ACTTTGACTGGAGAGAGTGAGATCCCTGATGTTTCTCAGTGCCCATTTGATGATGAAGA 240
Qy 297 ACCGAGATCCA----- 308
241 ACCGAGATCCA----- 300
309 ----- 308
301 CCAACACAAATCTTTTCTCCGCTGGGATATTCGGATGGGCTACTTTACATCACACTCT 360
```

```
Qy 309 -----TGTTCTCTGATGACGTGCAAAACCCCTATATATATGGCCCTCAGTATATCAAGT 362
Db 361 GCATAGTGTCTCTGATGACGTGCAAAACCCCTATATATATGGCCCTCAGTATATCAAGT 420
Qy 363 ACTTCAATGATAAACCATTGATGAGGAACCTAGAACGGGACAAAGAGGTCTACTTGGATTG 422
Db 421 ACTTCAATGATAAACCATTGATGAGGAACCTAGAACGGGACAAAGAGGTCTACTTGGATTG 480
Qy 423 TGGAGTCTTTGGCCAAATTTGATGACTGCCAAATCATTTGCCCTATCTATGCTGACC 482
Db 481 TGGAGTCTTTGGCCAAATTTGATGACTGCCAAATCATTTGCCCTATCTATGCTGACC 540
Qy 483 TCTCCCTTAAATACAACTGTACAGGGCTAAATTTTGGGAAGGTGATGTTGGACGCTATA 542
Db 541 TCTCCCTTAAATACAACTGTACAGGGCTAAATTTTGGGAAGGTGATGTTGGACGCTATA 600
Qy 543 CTGATGTTAGTACGCGGTACAAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCC 602
Db 601 CTGATGTTAGTACGCGGTACAAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCC 660
Qy 603 TGATCTCTTCCAAAGTGGCAAGGAGCAATCGCGGCGCCACAGATTGACAAGAAAGGAC 662
Db 661 TGATCTCTTCCAAAGTGGCAAGGAGCAATCGCGGCGCCACAGATTGACAAGAAAGGAC 720
Qy 663 GGGCTGTCTCATGGACCTTCTCTGAGGAGAAATGTCATCCGAGAAATTTAAATGAGC 722
Db 721 GGGCTGTCTCATGGACCTTCTCTGAGGAGAAATGTCATCCGAGAAATTTAAATGAGC 780
Qy 723 TATACACGCGGCGCCAAAGAACTATCAAAAGGCTGGAGACAATATCCCTCAGGAGCAGCTG 782
Db 781 TATACACGCGGCGCCAAAGAACTATCAAAAGGCTGGAGACAATATCCCTCAGGAGCAGCTG 840
Qy 783 TGGCTTCAACCCCGCCACACAGTGTGATGGGGAACAAAGAGGATAAATAGATCTCTC 842
Db 841 TGGCTTCAACCCCGCCACACAGTGTGATGGGGAACAAAGAGGATAAATAGATCTCTC 900
Qy 843 ACTTTGGCAGTGTCTCTCTCTCTCAATTCAGGCTCTTTCCATACCCACCAAGCCTGAG 902
Db 901 ACTTTGGCAGTGTCTCTCTCTCTCAATTCAGGCTCTTTCCATACCCACCAAGCCTGAG 960
Qy 903 GCTGAGCCTTTTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 962
Db 961 GCTGAGCCTTTTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Qy 963 TCTGATTTTAAAGAGGACATCTAGGGAATGTCAGGACCCCTACAGGAAGCCCTGCCATGC 1022
1021 TCTGATTTTAAAGAGGACATCTAGGGAATGTCAGGACCCCTACAGGAAGCCCTGCCATGC 1080
Qy 1023 TGTGGCCCAACTGTTTCACTGGGAGCAAGAGAGATCTCATAGGAGGAGGGAATGCT 1082
Db 1081 TGTGGCCCAACTGTTTCACTGGGAGCAAGAGAGATCTCATAGGAGGAGGGAATGCT 1140
Qy 1083 TTCCCTCCAAAGCTGGGTCAGTGTGTTAACTGCTTATCAGCTATTTCAGACATCTCCATGG 1142
Db 1141 TTCCCTCCAAAGCTGGGTCAGTGTGTTAACTGCTTATCAGCTATTTCAGACATCTCCATGG 1200
Qy 1143 TTTTCCCATGAACCTCTGTGGTTTTCATCATCTCTCTTTAGTTGAACTGCACAGCTGGTT 1202
Db 1201 TTTTCCCATGAACCTCTGTGGTTTTCATCATCTCTCTTTAGTTGAACTGCACAGCTGGTT 1260
Qy 1203 AGACCTAGATTTAAACCTTAAGTAAGATGCTGGGTATAGAACGCTAAGAAATTTTCCCTCC 1262
Db 1261 AGACCTAGATTTAAACCTTAAGTAAGATGCTGGGTATAGAACGCTAAGAAATTTTCCCTCC 1320
Qy 1263 AAGGACTCTTGTCTTCTTAAAGCCCTTCTGGCTTCTGTTTATGCTTCTTAAAGATATAA 1322
Db 1321 AAGGACTCTTGTCTTCTTAAAGCCCTTCTGGCTTCTGTTTATGCTTCTTAAAGATATAA 1380
Qy 1323 GCCTAACTTTGTGCTAGTCTCTTAAGGAGAAACCTTTAAACCAACAAAGTTTTTATCATTTGA 1382
Db 1381 GCCTAACTTTGTGCTAGTCTCTTAAGGAGAAACCTTTAAACCAACAAAGTTTTTATCATTTGA 1440
```

QY 1383 GACAATATTGAACACCCCTATTCTGGGGATTGAGAAGGGGTGAATAGAGGCTTGAG 1442
|||||
Db 1441 GACAATATTGAACACCCCTATTCTGGGGATTGAGAAGGGGTGAATAGAGGCTTGAG 1500
|||||
QY 1443 ACTTTCCTTTGTGTAGGACTTGGAGGAGAAATCCCTGGACTTTCACCTAACCCCTCTG 1502
|||||
Db 1501 ACTTTCCTTTGTGTAGGACTTGGAGGAGAAATCCCTGGACTTTCACCTAACCCCTCTG 1560
|||||
QY 1503 ACATACTCCCCACACCCAGTTGATGGCTTTCCTAATAAAAAAGATTGGGATT 1554
|||||
Db 1561 ACATACTCCCCACACCCAGTTGATGGCTTTCCTAATAAAAAAGATTGGGATT 1612
|||||

Search completed: July 12, 2003, 11:10:04
Job time : 218 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:54:20 ; Search time 23 Seconds
(without alignments)
1305.926 Million cell updates/sec

Title: US-09-954-846-2
Perfect score: 1359
Sequence: 1 MAVLAPLIALVYSPRLSRW.....EQVASTPTTVSDGKNDK 258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1359	100.0	258	10	US-09-954-846-2 Sequence 2, Appl
2	1359	100.0	272	10	US-09-925-300-1691 Sequence 1691, Ap
3	1326	97.6	296	9	US-09-905-291A-207 Sequence 207, App
4	1326	97.6	296	9	US-09-902-853-207 Sequence 207, App
5	1326	97.6	296	9	US-09-907-824-207 Sequence 207, App
6	1326	97.6	296	9	US-09-907-841-207 Sequence 207, App
7	1326	97.6	296	9	US-09-904-011-207 Sequence 207, App
8	1326	97.6	296	9	US-10-174-590-32 Sequence 32, Appl
9	1326	97.6	296	9	US-10-176-758-32 Sequence 32, Appl
10	1326	97.6	296	9	US-10-175-737-32 Sequence 32, Appl
11	1326	97.6	296	9	US-09-906-742-207 Sequence 207, App
12	1326	97.6	296	9	US-10-173-706-32 Sequence 32, Appl
13	1326	97.6	296	9	US-10-175-738-32 Sequence 32, Appl
14	1326	97.6	296	9	US-10-175-752-32 Sequence 32, Appl
15	1326	97.6	296	9	US-10-176-482-32 Sequence 32, Appl
16	1326	97.6	296	9	US-10-176-757-32 Sequence 32, Appl
17	1326	97.6	296	9	US-10-176-913-32 Sequence 32, Appl
18	1326	97.6	296	9	US-10-180-552-32 Sequence 32, Appl
19	1326	97.6	296	9	US-10-180-557-32 Sequence 32, Appl

20	1326	97.6	296	9	US-09-906-838-207 Sequence 207, App
21	1326	97.6	296	9	US-09-907-613-207 Sequence 207, App
22	1326	97.6	296	9	US-09-907-942-207 Sequence 207, App
23	1326	97.6	296	9	US-10-173-700-32 Sequence 32, Appl
24	1326	97.6	296	9	US-10-174-572-32 Sequence 32, Appl
25	1326	97.6	296	9	US-10-174-579-32 Sequence 32, Appl
26	1326	97.6	296	9	US-10-174-582-32 Sequence 32, Appl
27	1326	97.6	296	9	US-10-174-588-32 Sequence 32, Appl
28	1326	97.6	296	9	US-10-175-739-32 Sequence 32, Appl
29	1326	97.6	296	9	US-10-175-740-32 Sequence 32, Appl
30	1326	97.6	296	9	US-10-175-743-32 Sequence 32, Appl
31	1326	97.6	296	9	US-10-176-488-32 Sequence 32, Appl
32	1326	97.6	296	9	US-10-176-492-32 Sequence 32, Appl
33	1326	97.6	296	9	US-10-176-747-32 Sequence 32, Appl
34	1326	97.6	296	9	US-10-176-750-32 Sequence 32, Appl
35	1326	97.6	296	9	US-10-176-985-32 Sequence 32, Appl
36	1326	97.6	296	9	US-10-176-987-32 Sequence 32, Appl
37	1326	97.6	296	9	US-10-176-991-32 Sequence 32, Appl
38	1326	97.6	296	9	US-10-176-992-32 Sequence 32, Appl
39	1326	97.6	296	9	US-10-176-993-32 Sequence 32, Appl
40	1326	97.6	296	9	US-10-184-658-32 Sequence 32, Appl
41	1326	97.6	296	9	US-10-173-695-32 Sequence 32, Appl
42	1326	97.6	296	9	US-10-173-697-32 Sequence 32, Appl
43	1326	97.6	296	9	US-10-173-705-32 Sequence 32, Appl
44	1326	97.6	296	9	US-10-174-576-32 Sequence 32, Appl
45	1326	97.6	296	9	US-10-174-585-32 Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-09-954-846-2
; Sequence 2, Application US/09954846
; Patent No. US20020102654A1
; GENERAL INFORMATION:
; APPLICANT: Tanq, Y. Tom
; Guegler, Neil C.
; Patterson, Chandra
; Baughn, Mariah R.
; TITLE OF INVENTION: THIOREDOXIN PROTEINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/954,846
; FILING DATE: 17-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/107,248
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Carrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0556 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-855-0572
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 258 amino acids

```
;
;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT19
; CLONE: 324141
; SEQUENCE DESCRIPTION: SEQ ID NO: 2 :
US-09-954-846-2

Query Match      100.0%; Score 1359; DB 10; Length 258;
Best Local Similarity 100.0%; Pred. No. 3.9e-133;
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVLAPLIALVYSPRLSRWLAOPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNPCD 60
Db 1 MAVLAPLIALVYSPRLSRWLAOPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNPCD 60

Qy 61 FDWREVEILMFLSAIVMKNRRSMFLMTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTV 120
Db 61 FDWREVEILMFLSAIVMKNRRSMFLMTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTV 120

Qy 121 IVEFFANWSDCOSFAPYADLSLKYNCTGLNFGKVDVGRYTDVSTRYKYSTPLTKQLP 180
Db 121 IVEFFANWSDCOSFAPYADLSLKYNCTGLNFGKVDVGRYTDVSTRYKYSTPLTKQLP 180

Qy 181 TLILFOGKKEAMRRPQIDKKGRAVSWTFSEENVIREFNELNYORAKKLSKAGDNIPEEQ 240
Db 181 TLILFOGKKEAMRRPQIDKKGRAVSWTFSEENVIREFNELNYORAKKLSKAGDNIPEEQ 240

Qy 241 PVASTPTTVSDGENKKDK 258
Db 241 PVASTPTTVSDGENKKDK 258

RESULT 2
US-09-925-300-1691
; Sequence 1691, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL01
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1691
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1691

Query Match      100.0%; Score 1359; DB 10; Length 272;
Best Local Similarity 100.0%; Pred. No. 3.9e-133;
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVLAPLIALVYSPRLSRWLAOPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNPCD 60
Db 15 MAVLAPLIALVYSPRLSRWLAOPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNPCD 74

Qy 61 FDWREVEILMFLSAIVMKNRRSMFLMTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTV 120
Db 75 FDWREVEILMFLSAIVMKNRRSMFLMTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTV 134

Qy 121 IVEFFANWSDCOSFAPYADLSLKYNCTGLNFGKVDVGRYTDVSTRYKYSTPLTKQLP 180
Db 135 IVEFFANWSDCOSFAPYADLSLKYNCTGLNFGKVDVGRYTDVSTRYKYSTPLTKQLP 194

Qy 181 TLILFOGKKEAMRRPQIDKKGRAVSWTFSEENVIREFNELNYORAKKLSKAGDNIPEEQ 240
Db 181 TLILFOGKKEAMRRPQIDKKGRAVSWTFSEENVIREFNELNYORAKKLSKAGDNIPEEQ 240

Qy 241 PVASTPTTVSDGENKKDK 258
Db 241 PVASTPTTVSDGENKKDK 258

Qy 241 PVASTPTTVSDGENKKDK 258
Db 255 PVASTPTTVSDGENKKDK 272

RESULT 3
US-09-905-291A-207
; Sequence 207, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pah, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
```


; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 207
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-207

Query Match 97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

1 MAVLAPLIALVSVPRLSRWLAQPYLLSALLSAAFLLVKRLPLCHGLPTQREDGNPCD 60
|||||
1 MAVLAPLIALVSVPRLSRWLAQPYLLSALLSAAFLLVKRLPLCHGLPTQREDGNPCD 60
61 FQWREVEILMFLSAIVMKNRRS----- 83
61 FQWREVEILMFLSAIVMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
84 -MFLMTCCKPLYMGPEYIKYFNDKTIIDEELERDKRVTWIVEFFANWSNDCOSFAPYADL 142
121 IVFLMTCCKPLYMGPEYIKYFNDKTIIDEELERDKRVTWIVEFFANWSNDCOSFAPYADL 180
143 SLKYNTCTGLNFGKVDVGRYTDVSTRYKVSPTLKQPLTLILFOGGKEAMRRPQIDKKGR 202
181 SLKYNTCTGLNFGKVDVGRYTDVSTRYKVSPTLKQPLTLILFOGGKEAMRRPQIDKKGR 240
203 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
241 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296

RESULT 4

US-09-902-853-207
; Sequence 207, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350

; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 207
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-853-207

Query Match 97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

Qy 1 MAVLAPLIALVSVPRLSRWLAQPYLLSALLSAAFLLVKRLPLCHGLPTQREDGNPCD 60
Db 1 MAVLAPLIALVSVPRLSRWLAQPYLLSALLSAAFLLVKRLPLCHGLPTQREDGNPCD 60
Qy 61 FQWREVEILMFLSAIVMKNRRS----- 83
Db 61 FQWREVEILMFLSAIVMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
Qy 84 -MFLMTCCKPLYMGPEYIKYFNDKTIIDEELERDKRVTWIVEFFANWSNDCOSFAPYADL 142
Db 121 IVFLMTCCKPLYMGPEYIKYFNDKTIIDEELERDKRVTWIVEFFANWSNDCOSFAPYADL 180
Qy 143 SLKYNTCTGLNFGKVDVGRYTDVSTRYKVSPTLKQPLTLILFOGGKEAMRRPQIDKKGR 202
Db 181 SLKYNTCTGLNFGKVDVGRYTDVSTRYKVSPTLKQPLTLILFOGGKEAMRRPQIDKKGR 240
Qy 203 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
Db 241 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296

RESULT 5
US-09-907-824-207
; Sequence 207, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David

;; PRIOR APPLICATION NUMBER: PCT/US99/23089
;; PRIOR FILING DATE: 1999-10-05
;; PRIOR APPLICATION NUMBER: PCT/US99/28214
;; PRIOR FILING DATE: 1999-11-29
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 423
;; SEQ ID NO 207
;; LENGTH: 296
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-907-841-207

Query Match 97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

1 MAVLAPLALVSVPRLSRWLAOPYLLSALLSAAFLLVKRLPPLCHGLPTQREDGNPCD 60
|||||
1 MAVLAPLALVSVPRLSRWLAOPYLLSALLSAAFLLVKRLPPLCHGLPTQREDGNPCD 60
|||||

61 FDMREVEILMFLSAIVMKNNRS----- 83
|||||

61 FDMREVEILMFLSAIVMKNNRSITVEQHIIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
|||||

84 -MFLMTCPPLYMGPEYIKYFNDKTIIDELEDRKRVTWIVEFFANWSNDCQSFAPYADL 142
|||||

121 IVFLMTCPPLYMGPEYIKYFNDKTIIDELEDRKRVTWIVEFFANWSNDCQSFAPYADL 180
|||||

143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTPLTKOLPTLILFOGCKEAMRRPQIDKKGR 202
|||||

181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTPLTKOLPTLILFOGCKEAMRRPQIDKKGR 240
|||||

203 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
|||||

241 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296
|||||

RESULT 7
US-09-904-011-207
; Sequence 207, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350

;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: PCT/US00/04414
;; PRIOR FILING DATE: 2000-02-22
;; PRIOR APPLICATION NUMBER: US 60/143,048
;; PRIOR FILING DATE: 1999-07-07
;; PRIOR APPLICATION NUMBER: US 60/145,698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: US 60/146,222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: PCT/US99/20594
;; PRIOR FILING DATE: 1999-09-08
;; PRIOR APPLICATION NUMBER: PCT/US99/20944
;; PRIOR FILING DATE: 1999-09-13
;; PRIOR APPLICATION NUMBER: PCT/US99/21090
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/21547
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/23089
;; PRIOR FILING DATE: 1999-10-05
;; PRIOR APPLICATION NUMBER: PCT/US99/28214
;; PRIOR FILING DATE: 1999-11-29
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: 1999-11-30
;; PRIOR APPLICATION NUMBER: PCT/US99/28564
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/28565
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/30095
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: PCT/US99/30911
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US99/30999
;; PRIOR FILING DATE: 1999-12-20
;; PRIOR APPLICATION NUMBER: PCT/US00/00219
;; PRIOR FILING DATE: 2000-01-05
;; NUMBER OF SEQ ID NOS: 423
;; SEQ ID NO 207
;; LENGTH: 296
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-09-904-011-207

Query Match 97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

QY 1 MAVLAPLALVSVPRLSRWLAOPYLLSALLSAAFLLVKRLPPLCHGLPTQREDGNPCD 60
|||||
DB 1 MAVLAPLALVSVPRLSRWLAOPYLLSALLSAAFLLVKRLPPLCHGLPTQREDGNPCD 60
|||||

QY 61 FDMREVEILMFLSAIVMKNNRS----- 83
|||||

DB 61 FDMREVEILMFLSAIVMKNNRSITVEQHIIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
|||||

QY 84 -MFLMTCPPLYMGPEYIKYFNDKTIIDELEDRKRVTWIVEFFANWSNDCQSFAPYADL 142
|||||

DB 121 IVFLMTCPPLYMGPEYIKYFNDKTIIDELEDRKRVTWIVEFFANWSNDCQSFAPYADL 180
|||||

QY 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTPLTKOLPTLILFOGCKEAMRRPQIDKKGR 202
|||||

DB 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTPLTKOLPTLILFOGCKEAMRRPQIDKKGR 240
|||||

QY 203 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
|||||

DB 241 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296
|||||

RESULT 8
US-10-174-590-32
; Sequence 32, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; SEQ ID NO 32
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-32

Query Match          97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

QY 1 MAVLAPLIALVYSPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNCPCD 60
Db 1 MAVLAPLIALVYSPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNCPCD 60
QY 61 FDWREVEILMFLSAIVMMKNRRS----- 83
Db 61 FDWREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
QY 84 -MFLMTCKPPLYMGPEYIKYFNDKTIDELEDRKRVTWIVFFANWSNDCOSFAPFIADL 142
Db 121 IVFLMTCKPPLYMGPEYIKYFNDKTIDELEDRKRVTWIVFFANWSNDCOSFAPFIADL 180
QY 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 202
Db 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 240
QY 203 AVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
Db 241 AVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296

RESULT 9
US-10-176-758-32
; Sequence 32, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior application removed - See File Wrapper or Palm
; SEQ ID NO 32
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-32

Query Match          97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

QY 1 MAVLAPLIALVYSPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNCPCD 60
Db 1 MAVLAPLIALVYSPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNCPCD 60
QY 61 FDWREVEILMFLSAIVMMKNRRS----- 83
Db 61 FDWREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
QY 84 -MFLMTCKPPLYMGPEYIKYFNDKTIDELEDRKRVTWIVFFANWSNDCOSFAPFIADL 142
Db 121 IVFLMTCKPPLYMGPEYIKYFNDKTIDELEDRKRVTWIVFFANWSNDCOSFAPFIADL 180
QY 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 202
Db 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 240
QY 203 AVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
Db 241 AVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296

RESULT 9
US-10-176-758-32
; Sequence 32, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior application removed - See File Wrapper or Palm
; SEQ ID NO 32
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-32

Query Match          97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

QY 1 MAVLAPLIALVYSPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNCPCD 60
Db 1 MAVLAPLIALVYSPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNCPCD 60
QY 61 FDWREVEILMFLSAIVMMKNRRS----- 83
Db 61 FDWREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
QY 84 -MFLMTCKPPLYMGPEYIKYFNDKTIDELEDRKRVTWIVFFANWSNDCOSFAPFIADL 142
Db 121 IVFLMTCKPPLYMGPEYIKYFNDKTIDELEDRKRVTWIVFFANWSNDCOSFAPFIADL 180
QY 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 202
Db 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 240
QY 203 AVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
Db 241 AVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296
```

```
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-32

Query Match          97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

QY 1 MAVLAPLIALVYSPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNCPCD 60
Db 1 MAVLAPLIALVYSPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNCPCD 60
QY 61 FDWREVEILMFLSAIVMMKNRRS----- 83
Db 61 FDWREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
QY 84 -MFLMTCKPPLYMGPEYIKYFNDKTIDELEDRKRVTWIVFFANWSNDCOSFAPFIADL 142
Db 121 IVFLMTCKPPLYMGPEYIKYFNDKTIDELEDRKRVTWIVFFANWSNDCOSFAPFIADL 180
QY 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 202
Db 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 240
QY 203 AVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
Db 241 AVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296

RESULT 10
US-10-175-737-32
; Sequence 32, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 32
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-32

Query Match          97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

QY 1 MAVLAPLIALVYSPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNCPCD 60
Db 1 MAVLAPLIALVYSPRLSRWLAQPYLLSALLSAAFLVLRKLPPLCHGLPTQREDGNCPCD 60
QY 61 FDWREVEILMFLSAIVMMKNRRS----- 83
Db 61 FDWREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
QY 84 -MFLMTCKPPLYMGPEYIKYFNDKTIDELEDRKRVTWIVFFANWSNDCOSFAPFIADL 142
Db 121 IVFLMTCKPPLYMGPEYIKYFNDKTIDELEDRKRVTWIVFFANWSNDCOSFAPFIADL 180
QY 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 202
Db 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 240
QY 203 AVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
Db 241 AVSWTFSEENVIREFNLNELYQRAKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296
```

QY 143 SLKYNTGLNFGKVDVGRYTDVSTRYKYSTSLTKOLPTLLIFQGGKEAMRRPQIDKKGR 202
DB 181 SLKYNTGLNFGKVDVGRYTDVSTRYKYSTSLTKOLPTLLIFQGGKEAMRRPQIDKKGR 240
QY 203 AVSWTFSEENVIREFNELNYORAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
DB 241 AVSWTFSEENVIREFNELNYORAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296

RESULT 11

US-09-906-742-207
; Sequence 207, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,742
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 207
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-742-207

Query Match 97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;
QY 1 MAVLAPLIAIVSVPRLSRWLAOPYLLSALLSAAFLVVRKLPPLCHGLPTQREDGNPCD 60
DB 1 MAVLAPLIAIVSVPRLSRWLAOPYLLSALLSAAFLVVRKLPPLCHGLPTQREDGNPCD 60
QY 61 FDMREVEILMFLSAIVMKNNRS----- 83
DB 61 FDMREVEILMFLSAIVMKNNRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
QY 84 -MFLMTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANMSNDCQSFAPYADL 142
DB 121 IVFLMTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANMSNDCQSFAPYADL 180
QY 143 SLKYNTGLNFGKVDVGRYTDVSTRYKYSTSLTKOLPTLLIFQGGKEAMRRPQIDKKGR 202
DB 181 SLKYNTGLNFGKVDVGRYTDVSTRYKYSTSLTKOLPTLLIFQGGKEAMRRPQIDKKGR 240
QY 203 AVSWTFSEENVIREFNELNYORAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
DB 241 AVSWTFSEENVIREFNELNYORAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296

RESULT 12

US-10-173-706-32
; Sequence 32, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 32
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-706-32

Query Match 97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

```
QY 1 MAVLAPLIALVSVPRLSRWLAQPYLLSALLSAAFLVRLKLPPLCHGLPTOREDGNPCD 60
Db 1 MAVLAPLIALVSVPRLSRWLAQPYLLSALLSAAFLVRLKLPPLCHGLPTOREDGNPCD 60
QY 61 FDWREVEILMFLSAIVMMKNRRS----- 83
Db 61 FDWREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
QY 84 -MFLMTCKPPLYMGPEYIKYFNNDKTIDELERDKRVTWIVEFFANWSNDCQSFAPYADL 142
Db 121 IVFLMTCKPPLYMGPEYIKYFNNDKTIDELERDKRVTWIVEFFANWSNDCQSFAPYADL 180
QY 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 202
Db 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 240
QY 203 AVSWTFSEENVIREFNELNYORAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
Db 241 AVSWTFSEENVIREFNELNYORAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296
```

RESULT 13

---10-175-738-32

sequence 32, Application US/10175738

Publication No. US20030022294A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

ACIDS ENCODING THE SAME

FILE REFERENCE: P3430RIC45

CURRENT APPLICATION NUMBER: US/10/175,738

CURRENT FILING DATE: 2002-06-19

Prior application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 32

LENGTH: 296

TYPE: PRT

ORGANISM: Homo Sapien

US-10-175-738-32

Query Match 97.6%; Score 1326; DB 9; Length 296;

Best Local Similarity 86.8%; Pred. No. 1.2e-129;

Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

```
QY 1 MAVLAPLIALVSVPRLSRWLAQPYLLSALLSAAFLVRLKLPPLCHGLPTOREDGNPCD 60
Db 1 MAVLAPLIALVSVPRLSRWLAQPYLLSALLSAAFLVRLKLPPLCHGLPTOREDGNPCD 60
QY 61 FDWREVEILMFLSAIVMMKNRRS----- 83
Db 61 FDWREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
QY 84 -MFLMTCKPPLYMGPEYIKYFNNDKTIDELERDKRVTWIVEFFANWSNDCQSFAPYADL 142
Db 121 IVFLMTCKPPLYMGPEYIKYFNNDKTIDELERDKRVTWIVEFFANWSNDCQSFAPYADL 180
QY 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 202
Db 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKQLPTLLILFOGGKEAMRRPQIDKKGR 240
QY 203 AVSWTFSEENVIREFNELNYORAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 258
Db 241 AVSWTFSEENVIREFNELNYORAKKLSKAGDNIPPEQPVASTPTTVSDGENKKDK 296
```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176.482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 32
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-482-32

Query Match          97.6%; Score 1326; DB 9; Length 296;
Best Local Similarity 86.8%; Pred. No. 1.2e-129;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

      1 MAVLAPLIAIVSVPRLSRWLAQPYLLSALLSAAFLLVKRLPPLCHGLPTQREDGNPCD 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1 MAVLAPLIAIVSVPRLSRWLAQPYLLSALLSAAFLLVKRLPPLCHGLPTQREDGNPCD 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY      61 FDWREVEILMFLSAIVMMKNRRS----- 83
Db      61 FDWREVEILMFLSAIVMMKNRRSTVTEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
      84 -MFLMTCCKPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWNSDCOSFAPYIADL 142
      :||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 IVFLMTCCKPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWNSDCOSFAPYIADL 180
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY      143 SLKYNTCTGLNFGKVDVGRYTDVSTRYKVSTSPLTQKLTLLIFOGGKEAMRRPQIDKKGR 202
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 SLKYNTCTGLNFGKVDVGRYTDVSTRYKVSTSPLTQKLTLLIFOGGKEAMRRPQIDKKGR 240
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY      203 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPEQPVASTPTTVSDGENKKDK 258
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AVSWTFSEENVIREFNELNYQRAKKLSKAGDNIPEQPVASTPTTVSDGENKKDK 296

Search completed: July 9, 2003, 11:58:31
Job time : 25 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 10:39:22 ; Search time 1105 Seconds
(without alignments)
3781.390 Million cell updates/sec

Title: US-09-954-846-2
Perfect score: 1359
Sequence: 1 MAVLAPLIALVYSPRLSRW.....BQVASTPTTVSDGENKKDK 258

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Deiop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09954846/runat_09072003_102145_12167/app.query.fasta_1.455
-DB=EST -QEXT=FASTA -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.coi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09954846 -CGN_1_1456_runat_09072003_102145_12167 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

tabase :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_oth:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1354	99.6	1022	13	BM555220
2	1335	98.2	919	14	BQ961752
3	1318	97.0	884	14	BQ430136
4	1302	95.8	893	9	AL543426
5	1284	94.5	913	14	BQ920164
6	1265	93.1	1020	14	BM923588
7	1255.5	92.4	793	13	BI771888
8	1254	92.3	907	9	AL530852
9	1251	92.1	1097	14	BQ073145
10	1246.5	91.7	1030	14	BM905350
11	1243	91.5	875	14	BQ937601
12	1241	91.3	1076	14	BM921414
13	1231	90.6	938	14	BQ717825
14	1223	90.0	883	14	BQ439896
15	1220.5	89.8	1969	11	AK009759
16	1217.5	89.6	859	9	AL544734
17	1210	89.0	858	14	BQ230217
18	1209.5	89.0	890	9	AL515717
19	1199	88.2	724	12	BE741788
20	1183	87.0	1057	13	BM556247
21	1179	86.8	948	12	BE796785
22	1173	86.3	1006	13	BM562846
23	1171	86.2	1053	14	BQ070250
24	1170	86.1	933	14	BQ652172
25	1159	85.3	767	14	BM974314
26	1159	85.3	942	14	BQ957743
27	1154	84.9	991	14	BQ691103
28	1148	84.5	969	13	BM450430
29	1143	84.1	858	14	BQ686696
30	1139	83.8	900	13	BI838873
31	1136	83.6	937	13	BI598053
32	1132.5	83.3	883	13	BI335375
33	1126.5	82.9	939	14	BQ938032
34	1125	82.8	887	14	BQ687034
35	1125	82.8	894	13	BI654405
36	1118	82.3	811	13	BI550022
37	1115	82.0	830	13	BI764082
38	1109	81.6	763	9	AL545332
39	1108	81.5	797	13	BI521249
40	1108	81.5	905	13	BI836556
41	1091.5	80.3	892	14	BQ430854
42	1086	79.9	894	13	BI597723
43	1085.5	79.9	937	13	BI550817
44	1084	79.8	858	12	BG831608
45	1077	79.2	865	13	BI253341

ALIGNMENTS

RESULT 1
BM555220
LOCUS
DEFINITION AGENCOURT_6541254 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5550270
5', mRNA sequence.
ACCESSION BM555220
VERSION BM555220.1 GI:18795504
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1022)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12262 row: c column: 07
High quality sequence stop: 726.

FEATURES

source
1. 1022
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5550270"
/clone_lib="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
oligo-dr primed. Average insert size 1.767 Kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 246 a 261 c 255 g 251 t 9 others
ORIGIN
Alignment Scores:
Pred. No.: 1.72e-159 Length: 1022
Score: 1354.00 Matches: 257
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.63% Indels: 0
DB: 13 Gaps: 0
US-09-954-846-2 (1-258) x BM555220 (1-1022)

Qy 2 AlavalLeuAlaProLeuLeuAlaLeuValTyrsValProArGLeuSerArgTrpLeu 21
Db 1 GCGGTCTGGGCACCTCTAAATTCCTGCTGTCGTCGCGCGGACTTTCACGATGGCTC 60
Qy 22 AlaGlnProTyTrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuValArgLys 41
Db 61 GCCCAACCTTACTACTTCTGCGGCGCTCTCTGCTGCTTCTCTACTCTGTGAGGAA 120
Qy 42 LeuProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAspPhe 61
Db 121 CTGCGCGCGCTCTGCCAGGTCCTGCCACCAACCGGAAGACGGTAACCGTGTGACTTT 180
Qy 62 AsprTrpArgGluValGluLeuMetPheLeuSerAlaAlaLeuValMetMetLysAsnArg 81
Db 181 GACTGGAGAGAGTGGAGATCTGATGTTCTCAGTGCCTATGTGATGAAGAACCGC 240
Qy 82 ArgSerMetPheLeuMetThrCysLysProProLeuTyTrMetGlyProGluTyTrileLys 101
Db 241 AGATCCATGTTCTGTGATGACGTGCAACCCCTATATATGCGCCCTGATATATCAAG 300
Qy 102 TyrPheAsnAspLysThrIleAspGluLeuGluArgAspLysArgValThrTrpIle 121
Db 301 TACTTCAATGATAAACCATTGATGAGGAAGTACGAGGAGGAGGTCCTTGGATT 360
Qy 122 ValGluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyTrAlaAsp 141
Db 361 GTGGAGTCTTGTGCAATTTGGTCTTAATGACTGCCAATCATTTGCCCTTATATGCTGAC 420
Qy 142 LeuSerLeuLysTyTrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyTr 161
Db 421 CTCTCCCTTAATACAACTGTACAGGGCTAAATTTGGGAAGGTGGATTTGGAGCGCAT 480
Qy 162 ThrAspValSerThrArgTyTrLysValSerThrSerProLeuThrLysGlnLeuProThr 181
Db 481 ACTGATCTTAGTCGGGTACAAAGTAGGACATCACCCCTCACCAACCACTCCCTACC 540

Qy 182 LeuileLeuPheGlnGlyGlyLysGluAlaMetArgArgProGlnIleAspLysGly 201
Db 541 CTGATCCCTGTTCCAAAGTGGCAAGGAGCAATCGCGGCCACAGATTGACGAAGA 600
Qy 202 ArgAlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGlu 221
Db 601 CGGGCTGCTCATGACCTTCTCTCAGGAGAAATGTGATCCGAGATTTAACTTAATGAG 660
Qy 222 LeuTyTrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnPro 241
Db 661 CTATACCAAGCGGCCCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCGCT 720
Qy 242 ValAlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
Db 721 GTGGCTTCAACCCCAACACAGTGTGATGGGGAACCAAGAGGATAA 771

RESULT 2

LOCUS BQ961752 919 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8820550 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:6379836
5', mRNA sequence.
ACCESSION BQ961752
VERSION BQ961752.1 GI:22377230
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 919)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ruben Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2567 row: h column: 13
High quality sequence stop: 727.
Location/Qualifiers
1. 919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6379836"
/clone_lib="NIH_MGC_42"
/tissue_type="epitheloid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dr priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8Kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. I"

FEATURES

source
BASE COUNT 223 a 238 c 222 g 228 t 8 others
ORIGIN
Alignment Scores:
Pred. No.: 3.63e-157 Length: 919
Score: 1335.00 Matches: 254
Percent Similarity: 98.83% Conservative: 0
Best Local Similarity: 98.83% Mismatches: 3
Query Match: 98.23% Indels: 0
DB: 14 Gaps: 0
US-09-954-846-2 (1-258) x BQ961752 (1-919)

Db 637 GGAGGGCTGCTCATGACCTTCTCTNGAGAGATGTGATCCGAGATTTAACTTAAT 696
 QY 221 GluLeuTyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGln 240
 Db 697 GAGCTATACGAGCGGGCAAGAACTATCAAGGCTGGGAGACATATCCCTGAGGAGCAG 756
 QY 241 ProValAlaSerThrProThrValSerAspGlyGluAsnLysLysAsp 257
 Db 757 CTTGGCTTCAACCCACCACCGCTGTCAGATGGGGAACAGAGGGAT 807

RESULT 4

AL543426 893 bp mRNA linear EST 16-FEB-2001
 AL543426 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1003YH21 5
 DEFINITION prime, mRNA sequence.

ACCESSION AL543426

VERSION AL543426.1 GI:12875904

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 893)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

Location/Qualifiers

1..893

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0D1003YH21"

/clone_lib="LTI_NFL006_PL2"

/tissue_type="placenta"

/note="vector: pCWSPORT 6; Site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCWSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 233 a 228 c 206 g 224 t

ORIGIN

Annotation Scores:

i. No.: 4.99e-153 Length: 893
 Score: 1302.00 Matches: 252
 Percent Similarity: 86.64% Conservative: 1
 Best Local Similarity: 86.30% Mismatches: 1
 Query Match: 95.81% Indels: 38
 DB: 9 Gaps: 1

US-09-954-846-2 (1-258) x AL543426 (1-893)

QY 5 AlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTyrLeuAlaGlnPro 24
 Db 3 GCACCTCTAATGCTCTCGTGATTTCGGTCCGGGAGCTTCACCATGGCTCGCCCACT 62
 QY 25 TyrTyrLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArgLysLeuPro 44
 Db 63 TACTACCTTCTGTGGCCCTGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122
 QY 45 LeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAspPheAspTrpArg 64
 Db 123 CTCTGCCACGGTCTGCCACCCACCGGTAACCGGTAAACCGGTGACTTTGACTGGAGA 182

QY 65 GluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsnArgArgSer --- 83
 Db 183 GAATGGAGATCCCTGATGTTTCTCAGTCCCATGTGATGATGAAGAACCGAGATCCATC 242
 QY 83 ----- 83
 Db 243 ACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAAAGTGGCCCAACAACATTTCT 302
 QY 84 -----MetPheLeu 86
 Db 303 TCTTCCCTTGGATATTTCGCATGGGCTACTTTACATCACACTCTGCATAGTGTCTCTG 362
 QY 87 MetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyrPheAsnAspLys 106
 Db 363 ATGACGTCCAAACCCCTATATATGGCCCTGAGTATATCAAGTACTTCAATGATAAA 422
 QY 107 ThrIleAspGluGluLeuGluArgAspLysArgValThrTrpIleValGluPhePheAla 126
 Db 423 ACCATTGATAGGAACACTAGACGGGACAAGAGGGTCACCTGGATTGTGGAGTTCTTTGCC 482
 QY 127 AsnTyrSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLysSerLeuLysTyr 146
 Db 483 AATTGGTCTATGACTGCAATCATTTGCCCTATCTATGCTGACTCTCCCTTAAATAC 542
 QY 147 AsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThrAspValSerThr 166
 Db 543 AACTGTACAGGGCTAAATTTGGGAAGTGGATGTTGGACGCTATCATGATCTTAGTAGC 602
 QY 167 ArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeuIleLeuPheGln 186
 Db 603 CGGTACAAGTGCAGCACATCACCCCTCACCAAGCAACTCCCTACCTGATCTCTGTCCA 662
 QY 187 GlyGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArgAlaValSerTrp 206
 Db 663 GGTGCAAGGAGCAATCGCGGCCACAGATTGACAAAGAGGACGGCTCTCATAGG 722
 QY 207 ThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeuTyrGlnArgAla 226
 Db 723 ACCTTCTCTGAGGAGAATGTGATCGAGAAATTTAACTTAATGAGCTATACACAGGGCC 782
 QY 227 LysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProValAlaSerThrPro 246
 Db 783 AAGAACTATCAAGGCTGGAGAMAATATCCCTGAGGAGACGCTGTGGCTTCAACCCCC 842
 QY 247 ThrThrValSerAspGlyGluAsnLysLysAspLys 258
 Db 843 ACCACAGTCTCAGATGGGGAACAGAGGATAAA 878

BO920164 913 bp mRNA linear EST 20-AUG-2002
 AGENCOURT_8918912 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383675
 5', mRNA sequence.

BO920164

BO920164.1 GI:22334862

EST.

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 913)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: DCTD/BTP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

QY 4 LeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrpLeuAlaGln 23
 DB 19 TTGGACACCTTAATATTACTGTGATATTCGGTGGCCGACATTCACGATGGCTCGCCCAA 78
 QY 24 ProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArgLysLeuPro 43
 DB 79 CCTTACTACCTTCTGTGCGCCCTCTCTCTGTGCTTCTCTACTCTGTAGGAACATGCGG 138
 QY 44 ProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAspPheAspTrp 63
 DB 139 CGGCTCTGCCACGGTCTGCCACCCCAACGAGACGGTAACCCGCTGACTTTGACTGG 198
 QY 64 ArgGluValGluIleLeuMetPheLeuSerAlaIleValMetLysAsnArgArgSer 83
 DB 199 AGAAGATGGAGATCTGATGTTTCTCAGTGCATTTGTGATGATGAAGAACCGCATCC 258
 QY 83 -----MetPhe 85
 DB 259 ATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCCAACAAT 318
 QY 84 -----MetPhe 85
 DB 319 CTTTCTTCCGCTTGGATATTCGCATGGCCCTACTTTACATCACACTCTGCATAGTGTTC 378
 DB 86 LeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyrPheAsnAsp 105
 DB 379 CTGATGACGTGAACACCCCTATATATGGCCCTGAGTATATCAAGTACTTCAATGAT 438
 QY 106 LysThrIleAspGluLeuGluArgAspLysArgValTrpIleValGluPhePhe 125
 DB 439 AAACCATTTGATGAGAACTAGACGGGACAGAGGGTCACITGGATTGGAGTCTT 498
 QY 126 AlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeuSerLeuLys 145
 DB 499 GCCAATTGGTCTTAATGACTGCAATCATTTGCCCTATCTATGCTGACCTCTCCCTAAA 558
 QY 146 TyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTrpThrAspValSer 165
 DB 559 TACAACTGTACAGGCTAAATTTGGGAAGGTGGATGGAGCTTACTGATGTTAGT 618
 QY 166 ThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeuLeuPhe 185
 DB 619 ACGCGGTACAAAGTGGACATACACCCCTCACCAGCAACTCCCTACCTGATCCTGTTC 678
 QY 186 GlnGlyGlyLysGluAlaMetArgArgProGlnIleAspLysGlyArgAlaValSer 205
 DB 679 CAAGTGGCAAGAGGCAATGCGGCGCCACAGATGACAAGAAAGGAGCGGCTGTCTCA 738
 QY 205 TrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeuTyrGlnArg 225
 DB 739 TGGACCTTCTGTAGGAGAAATCGATCCGAGAAATTAATTAATGAGCTATACCAAGCG 798
 DB 226 -AlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProValAlaSerTh 245
 DB 799 GGCCAAAGAACTATCAAGAGCTGGAGACAATATCCCTCGAGAGCAGCCTGTGGCTTCCAC 858
 QY 245 rProThrThrValSerAspGlyGluAsnLysLys 256
 DB 859 CCCCACCCAGTGTCTAGAGGGGGAAACAAAGAAG 892
 RESULT 7
 LOCUS B1771888
 DEFINITION 603055295F1 NTH_MGC_122 Homo sapiens cDNA clone IMAGE:5204927 5',
 mRNA sequence.
 ACCESSION B1771888
 VERSION B1771888.1 GI:15763466
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 793)

NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue procurement: Life Technologies, Inc.
 cDNA Library preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLML1513 row: m column: 24
 High quality sequence stop: 775.
 Location/Qualifiers
 1..793
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5204927"
 /clone_lib="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
 Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH_MGC Library."
 BASE COUNT 206 a 202 c 197 g 188 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3e-147 Length: 793
 Score: 1255.50 Matches: 254
 Percent Similarity: 98.45% Conservative: 0
 Best Local Similarity: 98.45% Mismatches: 3
 Query Match: 92.38% Indels: 4
 DB: 13 Gaps: 0
 US-09-954-846-2 (1-258) x B1771888 (1-793)
 QY 1 MetAlaValLeuAlaProLeuIleAlaLeuValTyrSerValPro-ArgLeuSerArgTr 20
 DB 17 ATGGCGGTCTTGGACACCTTAATTCGTCGTGATTCGGTGGCCGACATTCACGATG 76
 QY 20 pLeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValAr 40
 DB 77 GCTGCGCCCAACCTTACTACTCTCTGTCGCCCTGCTCTCTGTCGCTTCTCTACTCTGAG 136
 QY 40 gLysLeuProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAs 60
 DB 137 GAAACTGCGCGCTCTGCCACGGTCTGCCACCCCAACGAGACGGTAACCCGTGTGA 196
 QY 60 pPheAspTrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAs 80
 DB 197 CTTTGTACTGGAGAAAGTGGAGATCCTGATGTTTCTCAGTGCCATTTGTGATGATGAAGA 256
 QY 80 nArgArgSerMetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIl 100
 DB 257 CCGCAGATCCATGTTCTCTGATGACGTGCAACCCCTATATATGCGCCCTGAGTATAT 316
 QY 100 eLysTyrPheAsnAspLysThrIleAspGluGluLeuGluArgAspLysArgValThrTr 120
 DB 317 CAAGTACTTCAATGATAAAACCATTCATGAGGAACATAGAACGGGACACAGAGGCTACTTG 376
 QY 120 pIleValGluPheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAl 140
 DB 377 GATTGTGGAGTCTTTGGCAATTTGGTCTAATGACTGCCAATCATTTGCCCTATCTATGC 436
 QY 140 aAspLeuSerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyAr 160

```

|||||
437 TGACCTCTCCCTTAATACAACTGTACAGGGCTAAATTTTGGGAGGTGGATGTTGGAGC 496
Qy 160 gTyrThrAspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuPr 180
Db 497 CTATACCTGATGTTAGTCGGGTACAAAGTCAGCACATCACCCCTCCACCAAGCAACTCCC 556
Qy 180 oThrLeuLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLy 200
Db 557 TACCCTGATCTGTTCCTCAAGGTGGCAAGAGGCAATGCGGGGCA-CAGATTGACAAGAA 615
Qy 200 sGlyArgAlaValSerThrPheSerGluGluAsnValIleArgGluPheAsnLeuAs 220
Db 616 AGGAGGGCTGCTCATGAGCACTTCTCTGAGAGAGATGTGATCCGAGAAATTAACCTTAAA 675
Qy 220 nGluLeuTyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluG 240
Db 676 TGAGCTATACAGCGGGC-AAGAAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCA 734
Qy 240 nProValAlaSerThrProThrThrValSerAspGlyGluAsnLysLysAsp 257
Db 735 GCCTGTGGCTTCAACCCAC--ACAGTGTACAGATGGGGAAACAAGAAGGAT 784

RESULT 8
AL530852 907 bp mRNA linear EST 13-FEB-2001
LOCUS AL530852 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD008YG11 5
DEFINITION prime, mRNA sequence.
ACCESSION AL530852
VERSION AL530852.1 GI:12794345
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DD008YG11"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 232 a 231 c 208 g 228 t 8 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 5.66e-147 Length: 907
Score: 1254.00 Matches: 245
Percent Similarity: 84.48% Conservative: 0
Best Local Similarity: 84.48% Mismatches: 7
Query Match: 92.27% Indels: 38
DB: 9 Gaps: 2

```

```

US-09-954-846-2 (1-258) x AL530852 (1-907)
Qy 7 LeuIleAlaLeuValTyrSerValProArgLeuSerArgTyrTrpLeuAlaGlnProTyrTyr 26
Db 3 CTAATTCGTCTCGTGTATTCGGTGGGACTTTCCAGTATGCTGCCTCGCCCAACTTACTAC 62
Qy 27 LeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArgLysLeuProLeuCys 46
Db 63 CTTCTGTCTGGCCCTCTCTCTGCTGCTTCTTACTCTGCTGAGGNAACTGCCCGCGCTCTGC 122
Qy 47 HisGlyLeuProThrGlnArgGluAspGlyAsnProCysAspPheAspThrPargGluVal 66
Db 123 CACGCTCTGCCACCCACCGCAAGCAGTAAACCCGTGTGACTGTGACTGTGAGAGAGTG 182
Qy 67 GluIleLeuMetPheLeuSerAlaIleValMetMetLysAsnArgArgSer----- 83
Db 183 GAGATCCCTGATGTTTCTCAGTGCCATTTGTGATGATGATGAAGAACCGCAGATCCATCATG 242
Qy 84 -----MetPhe----- 85
Db 243 GAGCAACATATAGGCAACATTTTTCATGTTTAAAGTGGCCCAACACAATTTCTTTCTTC 302
Qy 86 -----LeuMetThr 88
Db 303 CCCTTGGATATTCGCATGGCCCTACTTTACATCACACTCTGCATAGCTGTWCCNGAGAGC 362
Qy 89 CysLysProProLeuTyrMetGlyProGluTyrIleLysTyrPheAsnAspLysThrIle 108
Db 363 TCGAAACCCCTATATATGATGGCCCTGAGTATATCAAGTACTTCAATGATATAAACCAT 422
Qy 109 AspGluGluLeuGluArgAspLysArgValThrTrpIleValGluPhePheAlaSerTrp 128
Db 423 GATGAGGAACATAGAACGGGACAGAGGTCACCTTGGATTGTGGAGTCTTTTCCCAATTGG 482
Qy 129 SerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeuSerLeuLysTyrAsnCys 148
Db 483 TCTATGACTGCCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAAATACAACTGT 542
Qy 149 ThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThrAspValSerThrArgTyr 168
Db 543 ACAGGGCTAAATTTTGGGAAGGTGGATGTGGACGCTATACTGATGTAGTACGCGGTAC 602
Qy 169 LysValSerThrSerProLeuThrLysGlnLeuProThrLeuIleLeuPheGlnGlyGly 188
Db 603 AAAGTGACGCMCATMACCCCTCACCAAGCAACTCCCTACCTCATCTGTTCCAAGGTGC 662
Qy 189 LysGluAlaMetArgArgProGlnIleAspLysLysGlyArgAlaValSerTrpThrPhe 208
Db 663 AAGGAGGCAATGCGGGCCACAGATTGACAGAAGAGGACGGGCTGTCTCATGGACCTTC 722
Qy 209 SerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeuTyrGlnArgAlaLysLys 228
Db 723 TCTGAGGAGAAATGTTATCCGAGAATTTAACTTAAATGAGCYATACCAGCGGCCCAAGAA 782
Qy 229 LeuSerLysAlaGlyAspAsnIleProGluGluGlnProValAlaSerThrProThrThr 248
Db 783 CTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACCA 842
Qy 249 ValSerAspGlyGluAsnLysLysAspLys 258
Db 843 GTGTGAGATGGGGAACAAGAAGGATAAA 872

```

```

RESULT 9
LOCUS B0073145
DEFINITION B0073145 1097 bp mRNA linear EST 02-APR-2002.
AGENCOURT_6817826 NIH_MCC_118 Homo sapiens cDNA clone IMAGE:5757493
5', mRNA sequence.
ACCESSION B0073145
VERSION B0073145.1 GI:19902191
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1097)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cspbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM12799 row: m column: 14
 High quality sequence stop: 637.
 Location/Qualifiers
 1..1097
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_118"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated, adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC Library."
 BASE COUNT 261 a 296 c 276 g 263 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.8e-146 Length: 1097
 Score: 1251.00 Matches: 246
 Percent Similarity: 97.63% Conservative: 1
 Best Local Similarity: 97.23% Mismatches: 3
 Query Match: 92.05% Indels: 3
 Gaps: 0
 14
 US-09-954-846-2 (1-258) x B0073145 (1-1097)
 QY 3 ValLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTirPleuAla 22
 DB 1 GTCTGGCACCTCTAATGTCTCTGCTCTCGGTGGCGGACATTCACGATGGCTGCC 60
 QY 23 GlnProTyrTyrLeuLeuSerAlaLeuSerAlaAlaPheLeuLeuValArgLysLeu 42
 DB 61 CAACCTTACTACCTTCGTGGCCCTGCTCTGCTGCTTCCCTGCTGAGGAACTG 120
 QY 43 ProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAspPheAsp 62
 DB 121 CCGCCGCTGCGCACGCTGCGCCACCCACCGGAGACCGGTAAACCCGCTGACTTGCAC 180
 QY 63 TrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsnArgArg 82
 DB 181 TGAGAGAACTGGAGATCTGTGTTCTCAGTGCCATTGTGATGATGAGAACCGGAGA 240
 QY 83 SerMetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102
 DB 241 TCCATGTTCTGTAGTACGTGCAAAACCCCTATATATGCGCCCTGAGTATATCAAGCAC 300
 QY 103 PheAsnAspLysThrIleAspGluLeuGluArgAspLysArgValThrTrpIleVal 122
 DB 301 TTCAATGATAAAACCATTTGATGAGAACTAGACGGGACAGAGGGTCACTTGGATTGTG 360
 QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
 DB 361 GAGTCTTTCCTCAATTTGGTCTAATGACTGCGCAATCTTTGGCCCTATCTATGCTGACCTC 420

QY 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr 162
 DB 421 TCCCTTAATAACAACCTGTACAGGGCTAAATTTTGGGAAGGTGGATGTGGACGCTACT 480
 QY 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
 DB 481 GATGTTAGTACCGGTACAAAGTACAGACATCACCCCTCACCAAGCAACTCCCTACCTG 540
 QY 183 IleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArg 202
 DB 541 ATCTGTTCCTCAAGGTGGCAAGGAGCAATCGCGGCCACAGATTGACAAAGAGCGG 600
 QY 203 AlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu 222
 DB 601 GCTGCTCATGGACCTTCTGTGAGGAGAAATGTGATCCGAGAAATTTAACTTAAATGAGCTA 660
 QY 223 TyrGlnArgAlaLysLysLeuSerLysAla-GlyAspAsnIle-ProGluGluGlnProV 242
 DB 661 TACCCGGGGCCCAAGAACTATCAAGGCTGGGAGACAAATATCTCTGAGGAGCAGCCTG 720
 QY 242 alala-SerThrProThrThrValSerAspGly 252
 DB 721 TGGCCTTCAACCCCCACCACAGTGTGACATGGG 753
 RESULT 10
 BM905350
 LOCUS AGENCOURT_6698724 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5558249
 DEFINITION 5', mRNA sequence.
 ACCESSION BM905350
 VERSION BM905350.1 GI:19355729
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1030)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cspbs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM12280 row: o column: 18
 High quality sequence stop: 711.
 Location/Qualifiers
 1..1030
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."
 BASE COUNT 254 a 274 c 245 g 255 t 2 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.03e-146 Length: 1030
 Score: 1246.50 Matches: 249
 Percent Similarity: 84.90% Conservative: 4
 Best Local Similarity: 83.56% Mismatches: 4
 Query Match: 91.72% Indels: 41


```

QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
Db 492 GAGTTCCTTTGCCAATTGGTCTAATGACTGCGCAATCATTTGCCCTATCTATGCTGACCTC 551
QY 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr 162
Db 552 TCCCTTAATACAACTGTACAGGCTAAATTTGGGAAGGTGGATGTTGGACGCTATACT 611
QY 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
Db 612 GATGTTAGTACGGGTACAAAGTGAGCACATCACCCCTCACCAGCAACTCCCTACCCCTG 671
QY 183 IleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArg 202
Db 672 ATCCTGTTCGAAGGTGGCAAGGAGGCAATGGCGGCCACAGATGTGCAAGAAAGGACGG 731
QY 203 AlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu 222
Db 732 GCTGTCTCATGGACCTTCTCTGAGGAGATGTGATCCGAGAAATTAATTAATGAGCTA 791
QY 223 TyrGlnArgAlaLys-LysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVa 242
Db 792 TACCAGCGGCCCAAGAAACTATCAAGGCTGGAGACAATATCCTGAGGAGCAGCCTGT 851
QY 242 lAlaSerThrProThrThrVal 249
Db 852 GGCTTCAACCCCCACACAGGTG 873

RESULT 12
LOCUS BM921414 1076 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6626480 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752847
5', mRNA sequence.
ACCESSION BM921414
VERSION BM921414.1 GI:19371793
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1076)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M12787 row: k column: 24
High quality sequence stop: 697.
Location/Qualifiers
1. 1076
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752847"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
FEATURES
source

```

```

BASE COUNT 257 a 296 c 251 g 270 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 3.17e-145 Length: 1076
Score: 1241.00 Matches: 253
Percent Similarity: 85.28% Conservative: 2
Best Local Similarity: 84.62% Mismatches: 1
Query Match: 91.32% Indels: 43
DB: 14 Gaps: 1
US-09-954-846-2 (1-258) x BM921414 (1-1076)
QY 1 MetAlaValLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrp 20
Db 18 ATGCGCGTCTTTGGGACCTCTAAATGCTCTCGTATTGCGGTGCGCGACTTTCACGATGG 77.
QY 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuValArg 40
Db 78 CTCGCCCAACCTTACTACCTTCTCTGCGCCCTGCTCTGCTGCTTCTCTACTCGTGAGG 137
QY 41 LysLeuProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
Db 138 AAAGTGGCGCGCTCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 197
QY 61 PheAspTrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsn 80
Db 198 TTTGACTGGAGAGAAGTGGAGATCTGATGTTTCTCAGTGCCTTGTGATGATGAAGAAC 257
QY 81 ArgArgSer----- 83
Db 258 CGCAGATCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGCC 317
QY 83 ----- 83
Db 318 AACACAATTCCTTTCTCCGCTTGGATATTCGCATGGCGCTTACATCACACTCTGC 377
QY 84 ---MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102
Db 378 ATAGTGTTCCTGTGATGACGTGCAACCCCTCTATATATGCGCCCTGAGTATATCAAGTAC 437
QY 103 PheAsnAspLysThrIleAspGluGluLeuGluArgAspLysArgValThrTripleVal 122
Db 438 TTCATGATATAAACCACTGATGAGAACTAGACGGGACAGAGGGTGCCTTGGATGTG 497
QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
Db 498 GAGTTCCTTTGCCAATTTGGTCTAATGACTGCGCAATCATTTGCCCTATCTATGCTGACCTC 557
QY 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr 162
Db 558 TCCCTTAATACAACTGTACAGGCTAAATTTGGGAAGGTGGATGTTGGACGCTATACT 617
QY 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
Db 618 GATGTTAGTACCGGTACAAAGTGAGCACATCACCCCTCACCAGCAACTCCCTACCCCTG 677
QY 183 IleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArg 202
Db 678 ATCCTGTTCGAAGGTGGCAAGGAGCAATGGCGGCCACAGATGTGCAAGAAAGGACGG 737
QY 203 AlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu 222
Db 738 GCTGTCTCATGGACCTTCTCTGAGGAGAAATGTGATCCGAGAATTAATTAATGAGCTA 797
QY 223 TyrGlnArgAla-LysLysLeuSer-LysAlaGlyAspAsnIle-ProGluGluGlnPro 241
Db 798 TACCAGCGGCCCAAGAAACTATCAAGGCTGGAGACAATATTCCCTTGGAGAGCAGCCT 857
QY 242 Val-AlaSerThrProThrThrValSerAsp-GlyGluAsnLysLys 256
Db 858 GTGGCTTCAACCCCCACACAGTGCAGATGGCGGGGAAACACAGAG 904

```

```

RESULT 13
B0717825
LOCUS      B0717825      938 bp      mRNA      linear      EST 16-JUL-2002
DEFINITION AGENCOURT_8229576 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6184859 5', mRNA sequence.
ACCESSION  B0717825
VERSION    B0717825.1 GI:21856722
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 938)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
IDENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13574 row: 1 column: 12
            High quality sequence stop: 684.
FEATURES   source
            1..938
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:6184859"
            /clone_lib="Lupski_dorsal_root_ganglion"
            /sex="male"
            /tissue_type="dorsal root ganglia"
            /dev_stage="adult, 36 yr"
            /lab_host="DH10B"
            /note="vector: pCMV-SPORT6 (Life Technologies); Site_1:
            Note1; Site_2: SalI; cDNA made by oligo-dr priming.
            Directionally cloned using the following adaptors:
            5'-TCGACCCACGCGTCG-3' and
            5'-GACTAGTCTCTAGATCGGAGCGCGCCCT(15)-3'. Size selected >
            1 kb for average insert length 1.7 kb. This is a primary
            library, non-amplified. Library constructed by Life
            Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
            College of Medicine) and is available through Life
            Technologies."
BASE COUNT 238 a 244 c 221 g 235 t
ORIGIN
Alignment Scores:
Pred. No.: 4,69e-144 Length: 938
Score: 1231.00 Matches: 250
Percent Similarity: 84.28% Conservative: 2
Best Local Similarity: 83.61% Mismatches: 4
Query Match: 90.58% Indels: 43
DB: 14 Gaps: 1
US-09-954-846-2 (1-258) x B0717825 (1-938)
QY 1 MetAlaValLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrp 20
Db 8 ATGGCGGTCTTGGACCTCTAATGCTCTGCTGTTATTCGGTGGCGGACTTTCACGATGG 67
QY 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuValArg 40
Db 68 CTGCGCCCAACTTACTACCTTCTGTCGGCCCTGCTCTCTGCTGCTTCTCTACTCGTGAGG 127
QY 41 LysLeuProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
Db 128 AAATGCGCGCGCTCTGCCACGGTCTGCCACCCACCGAAGCGGTAAACCCGTGTGAC 187
QY 61 PheAspTrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsn 80

```

High quality sequence stop: 616.
 Location/Qualifiers
 1. .883
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6069357"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="organ: testis; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 223 a 230 c 210 g 219 t 1 others
 ORIGIN

Alignment Scores:
 pred. No.: 4.37e-143 Length: 883
 Score: 1223.00 Matches: 245
 Percent Similarity: 84.83% Conservatives: 1
 Best local Similarity: 84.48% Mismatches: 3
 Query Match: 89.99% Indels: 41
 14 Gaps: 1

US-09-954-846-2 (1-258) x B0439896 (1-883)

Qy 4 LeuAlaProLeuLeuAlaLeuValTySerValProArgLeuSerArgTrpLeuAlaGln 23
 |||||
 Db 2 TTGGCACCTCTAATGCTCTGCTATTCGGTGGCGGACCTTCACGATGCTCGCCAA 61
 |||||

Qy 24 ProTyTrpLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArgLysLeuPro 43
 |||||
 Db 62 CCTTACTACCTTCTGTGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 121
 |||||

Qy 44 ProLeuCyHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAspPheAspPrp 63
 |||||
 Db 122 CCCTCTGCCACGCTCTGCCACCCACCGAAGCGGTACCCGTGTGACTTTGACGCG 181
 |||||

Qy 64 ArgGluValGluLeuLeuMetPheLeuSerAlaLeuValMetMetLysAsnArgArgSer 83
 |||||
 Db 182 AGAGAAGTGGAGATCCGTGATGTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGATCC 241
 |||||

Qy 83 ----- 83

Db 242 ATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAAACACAAAT 301
 |||||

Qy 84 -----MetPhe 85
 |||||

Db 302 CTTTCTTCCGTTGGATATTCGATGGCGCTACTTTACATCACACTCTGCATAGTGTTC 361
 |||||

Qy 86 LeuMetThrCysLysProProLeuTyMetGlyProGluTyrlleLysTyTrpPheAsnAsp 105
 |||||
 Db 362 CTGATGACGTGCACAAACCCCTATATATGCGCCTGAGTATATCAAGTACTTCAATGAT 421
 |||||

Qy 106 LysThrIleAspGluLeuGluArgAspLysArgValThrTrpIleValGluPhePhe 125
 |||||
 Db 422 AAACCAATTGATGAGGAACATAGAACGGGCAAGAGGTCCTTGGATTGGAGTCTCTT 481
 |||||

Qy 126 AlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyAlaAspLeuSerLeuLys 145
 |||||
 Db 482 GCCAATTTGGTCTAATGATGCCAATCATTTGCCCTATCATGCTGACCTCTCCCTTAA 541
 |||||

Qy 146 TyraAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyTrpAspValSer 165
 |||||
 Db 542 TACAACCTGTACAGGGCTAAATTTTGGGAAGTGGTGTGGACGCTACTACTGATGTTAGT 601
 |||||

Qy 166 ThrArgTyrlsValSerThrSerProLeuThrLysGlnLeuProThrLeuIleLeuPhe 185
 |||||
 Db 602 ACCGGTACAAAGTGGACACATCACCCCTCACCAAGCAACTCCCTACCCCTGATCCTGTC 661
 |||||

Qy 186 GlnGlyCysGluAlaMetArgArgProGlnIleAspLysLysGlyArgAlaValSer 205
 |||||

Db 662 CAAGGTGCAAGGAGCAATCGCGGCCACACGATTGACAGAAAGGACGGCTCTCTCA 721
 |||||

Qy 206 TrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeuTyTrpGlnArg 225
 |||||

Db 722 TGGACCTTCTCTGAGAGAAATGTGATCGGAGAAATTTAACTTAATAGCATATACACGGG 781
 |||||

Qy 226 -AlaLysLysLeuSerLysAla-GlyAspAsnIleProGluGluGlnProVal-AlaSer 244
 |||||

Db 782 GGCCAAAGAACTATCAAGGCTGGGAGACAATATCCCTTGAGAGCACCTGTGGGCTTCA 841
 |||||

Qy 245 ThrProThrValSerAspGly 252
 |||||

Db 842 ACCCCACACACGATCGGATGGG 865

RESULT 15
 AK009759
 LOCUS
 DEFINITION
 Mus musculus adult male tongue cDNA, RIKEN full-length enriched
 library, clone:2310042M24:related to SIMILARITY TO AA 31-55 OF A.
 THALIANA THIOREDOXINE, full insert sequence.
 ACCESSION
 AK009759
 VERSION
 AK009759.1 GI:12844749
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
 clone.lib-RIKEN full-length enriched mouse cDNA library
 clone:2310042M24.
 ORGANISM
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci,P. and Hayashizaki,Y.
 TITLE
 High-efficiency full-length cDNA cloning
 JOURNAL
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 99279253
 PUBMED
 10349636
 2
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL
 Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE
 20499374
 PUBMED
 11042159
 3
 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
 Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
 Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL
 Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE
 20530913
 PUBMED
 11076861
 4
 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
 Arakawa,T., Harada,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
 Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaoka,I.,
 Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
 Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
 Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
 Kuehl,P., Lewis,S., Matsuo,Y., Nikolaev,I., Pesole,G.,
 Quackenbush,J., Schriml,L.M., Staabli,F., Suzuki,R., Tomita,M.,
 Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
 Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
 Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
 Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
 Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
 Marchionni,L., Mashima,J., Mazzarelli,J., Mombarts,P., Nordone,P.,
 Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
 Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,

Qy 243 AlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
Db 847 GACCCCTGCTCCCACTACTGTGCGCAGATGGGGAACACAGAGGACAAA 894

Search completed: July 12, 2003, 12:03:35
Job time : 1112 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:49:35 ; Search time 80 seconds
(without alignments)
664.502 Million cell updates/sec

Title: US-09-954-846-2
Perfect score: 1359
Sequence: 1 MAVLAPLIALVSVPRLSRW.....EQVASTPTVSDGENKDK 258

Working table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1326	97.6	296	Q9Y320	Q9Y320 homo sapien
2	1263	92.9	372	Q9H3L1	Q9H3L1 homo sapien
3	1220.5	89.8	295	Q9D710	Q9D710 mus musculus
4	476.5	35.1	271	Q9V8W7	Q9V8W7 drosophila
5	391	28.8	265	Q18484	Q18484 caenorhabdi
6	173	12.7	253	Q9LYG6	Q9LYG6 arabidopsis
7	138.5	10.2	106	Q8T9N5	Q8T9N5 schistosoma
8	135.5	10.0	100	Q9JN64	Q9JN64 mycoplasma
9	131.5	9.7	100	Q9R6P9	Q9R6P9 mycoplasma
10	126.5	9.3	121	Q8TFM8	Q8TFM8 fusarium cu
11	125	9.2	135	Q97W14	Q97W14 sulfolobus
12	124.5	9.2	406	Q16961	Q16961 acanthamoeb
13	121.5	8.9	145	Q8ZBW4	Q8ZBW4 versinia pe
14	120.5	8.9	105	Q28138	Q28138 archaeoglob
15	120.5	8.9	134	Q28984	Q28984 archaeoglob
16	119	8.8	486	Q96JJ7	Q96JJ7 homo sapien

17	119	8.8	493	5	Q17770	Q17770 caenorhabdi
18	118.5	8.7	141	5	Q27142	Q27142 euplotis cr
19	118.5	8.7	304	5	Q9N357	Q9N357 caenorhabdi
20	118	8.7	105	16	Q8XMF0	Q8XMF0 clostridium
21	117	8.6	108	16	Q8Y056	Q8Y056 ralstonia s
22	117	8.6	223	16	Q8RAI5	Q8RAI5 thermoanaer
23	116.5	8.6	104	16	Q8XHX1	Q8XHX1 clostridium
24	116	8.5	606	5	Q9TW67	Q9TW67 caenorhabdi
25	115.5	8.5	359	10	P93358	P93358 nicotiana t
26	115	8.5	612	4	Q96K36	Q96K36 homo sapien
27	114.5	8.4	393	5	Q01492	Q01492 caenorhabdi
28	114	8.4	139	16	Q67747	Q67747 aquifex ae
29	112.5	8.3	496	5	Q25598	Q25598 onchocerca
30	112	8.2	105	16	Q8YVT1	Q8YVT1 anabaena sp
31	111.5	8.2	127	16	Q8XAT2	Q8XAT2 escherichia
32	111.5	8.2	597	10	Q9FF55	Q9FF55 arabidopsis
33	110.5	8.1	104	2	Q48985	Q48985 mycoplasma
34	110	8.1	98	5	Q25549	Q25549 naegleria f
35	110	8.1	107	5	Q17486	Q17486 echinococcu
36	110	8.1	142	5	Q9W022	Q9W022 drosophila
37	110	8.1	160	3	Q8TGI0	Q8TGI0 podospira a
38	110	8.1	196	10	Q9FUL2	Q9FUL2 prunus aviu
39	109.5	8.1	140	17	Q96YQ0	Q96YQ0 sulfolobus
40	108.5	8.0	114	10	Q8S3L3	Q8S3L3 populus tre
41	108.5	8.0	139	5	Q9VUG9	Q9VUG9 drosophila
42	108.5	8.0	324	4	Q9BVH9	Q9BVH9 homo sapien
43	108.5	8.0	363	4	Q8TCT2	Q8TCT2 homo sapien
44	107.5	7.9	106	5	Q9U8F3	Q9U8F3 schistosoma
45	107.5	7.9	144	16	Q9KLJ2	Q9KLJ2 vibrio chol

ALIGNMENTS

RESULT 1

- Q9Y320 PRELIMINARY; PRT; 296 AA.
- AC Q9Y320;
 - DT 01-NOV-1999 (TREMBLrel. 12, Created)
 - DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 - DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 - DE CGI-31 protein.
 - OS Homo sapiens (Human).
 - OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 - OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 - OX NCBI_TaxID=9606;
 - RN [1]
 - RP SEQUENCE FROM N.A.
 - RA Lin W.-C.;
 - RT "Comparative gene cloning: Identification of novel human genes with C. elegans proteome as template."
 - RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 - RN [2]
 - RP SEQUENCE FROM N.A.
 - RC TISSUE=EYE;
 - RA Strausberg R.;
 - RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 - DR EMBL; AFI32965; AAD27740.1; -
 - DR EMBL; BC000666; AAH00666.1; -
| DR InterPro; IPR001005; Myb_DNA_binding. | |
| DR InterPro; IPR000063; ThioRed. | |
| DR PROSITE; PS00037; MYB_1; UNKNOWN_1. | |
| SQ SEQUENCE 296 AA; 34037 MW; D646B16B3A5F0C6D CRC64; | |

Query Match 97.6%; Score 1326; DB 4; Length 296;
Best Local Similarity 86.8%; Pred. No. 3.8e-121;
Matches 257; Conservative 1; Mismatches 0; Indels 38; Gaps 1;

QY 1 MAVLAPLIALVSVPRLSRWLAQPYLLSALLSAFLVLRKLPPLCHGLPTQREDGNPCD 60
|||||
DB 1 MAVLAPLIALVSVPRLSRWLAQPYLLSALLSAFLVLRKLPPLCHGLPTQREDGNPCD 60
|||||

QY 61 FDWREVEILMFLSAIYMKNNRRS----- 83

```
Db 61 FDWREVEILMFLSAIVMKNNRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
Qy 84 -MFLMTCPPPLYMGPEYIKYFNNDKTIIDELERDKRVTWIVEFFANWSDCQSFAPYADL 142
Db 121 IVFLMTCPPPLYMGPEYIKYFNNDKTIIDELERDKRVTWIVEFFANWSDCQSFAPYADL 180
Qy 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKOLPTLLILFQGGKAMRRPQIDKKGR 202
Db 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKOLPTLLILFQGGKAMRRPQIDKKGR 240
Qy 203 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPEQVPASTPTTSDGENKDK 258
Db 241 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPEQVPASTPTTSDGENKDK 296

RESULT 2
Q9H3L1 ID Q9H3L1 PRELIMINARY; PRT; 372 AA.
AC Q9H3L1;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
% My009 protein.
% Homo sapiens (Human).
% Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Mao Y.M., Xie Y., Zhou Z.X., Ying K., Zheng Z.H.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF059753; AAG43122.1;
DR InterPro; IPR001005; Myb_DNA_Binding.
DR InterPro; IPR000063; ThioRed.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
SQ SEQUENCE 372 AA; 42477 MW; DE94F9E11B4D8C4C CRC64;

Query Match 92.9%; Score 1263; DB 4; Length 372;
Best Local Similarity 83.3%; Pred. NO. 7.2e-115;
Matches 245; Conservative 4; Mismatches 7; Indels 38; Gaps 1;

Qy 1 MAVLAPLIALVYSPRLSRWLQAOPYLLSALLSRAFLVRKLPPLCHGLPTQREDGNPCD 60
Db 1 MAVLAPLIALVYSPRLSRWLQAOPYLLSALLSRAFLVRKLPPLCHGLPTQREDGNPCD 60
Qy 61 FDWREVEILMFLSAIVMKNNRS----- 83
Db 61 FDWREVEILMFLSAIVMKNNRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC 120
~ 84 -MFLMTCPPPLYMGPEYIKYFNNDKTIIDELERDKRVTWIVEFFANWSDCQSFAPYADL 142
121 IVFLMTCPPPLYMGPEYIKYFNNDKTIIDELERDKRVTWIVEFFANWSDCQSFAPYADL 180
Qy 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKOLPTLLILFQGGKAMRRPQIDKKGR 202
Db 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKOLPTLLILFQGGKAMRRPQISKGR 240
Qy 203 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPEQVPASTPTTSDGENK 256
Db 241 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPEQVPASTPQCQMGKTRR 294

RESULT 3
Q9D710 ID Q9D710 PRELIMINARY; PRT; 295 AA.
AC Q9D710;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE 2310042M24Rik protein (RIKEN cDNA 2310042M24 gene).
GN 2310042M24Rik.
```

```
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONGUE;
RX STRAIN=C57BL/6J; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saïto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo T., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK009759; BAB26483.1;
DR EMBL; BC013544; AAH13544.1;
DR MGD; MGI:1914208; 2310042M24Rik.
DR InterPro; IPR001005; Myb_DNA_Binding.
DR InterPro; IPR000063; ThioRed.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
SQ SEQUENCE 295 AA; 33942 MW; 25B4477F7BB17E3C CRC64;

Query Match 89.8%; Score 1220.5; DB 11; Length 295;
Best Local Similarity 80.4%; Pred. No. 7.4e-111;
Matches 238; Conservative 9; Mismatches 10; Indels 39; Gaps 2;

Qy 1 MAVLAPLIALVYSPRLSRWLQAOPYLLSALLSRAFLVRKLPPLCHGLPTQREDGNPCD 60
Db 1 MAVLAPLIALVYSPRLSRWLQAOPYLLSALLSRAFLVRKLPPLCHGLPTQREDGNPCD 60
Qy 61 FDWREVEILMFLSAIVMKNNRS----- 83
Db 61 FDWREVEILMFLSAIVMKNNRSITVEQHIGNIFMFSKVANAILFFRLDIRMGLLYITLC 120
Qy 84 -MFLMTCPPPLYMGPEYIKYFNNDKTIIDELERDKRVTWIVEFFANWSDCQSFAPYADL 142
121 IVFLMTCPPPLYMGPEYIKYFNNDKTIIDELERDKRVTWIVEFFANWSDCQSFAPYADL 180
Qy 143 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKOLPTLLILFQGGKAMRRPQIDKKGR 202
Db 181 SLKYNCTGLNFGKVDVGRYTDVSTRYKYSTSPLTKOLPTLLILFQGGKAMRRPQIDKKGR 240
Qy 203 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPEQVPASTPTTSDGENKDK 258
Db 241 AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPEQVPASTPTTSDGENKDK 295

RESULT 4
Q9V8W7 ID Q9V8W7 PRELIMINARY; PRT; 271 AA.
AC Q9V8W7;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
```

```

DE CG11007 protein.
GN CG11007.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinauer S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003795; AAF57540.1;
DR FlyBase; FBgn0034455; CG11007.
DR InterPro; IPR000063; ThioRed.
SQ SEQUENCE 271 AA; 31322 MW; 1D7A50C3B3B19F CRC64;

Query Match 35.1%; Score 476.5; DB 5; Length 271;
Best Local Similarity 38.9%; Pred. No. 2.8e-38;
Matches 96; Conservative 45; Mismatches 65; Indels 41; Gaps 3;

QY 21 LAOPYLLSALLSAFLVRLKPLCHGLPTQREDGNPCDFDREVEILMFSAIVMKN 80
Db 10 LAKPYWVNILLATSYLLAKKTQICITRLFLAGEDACDLDSREVEILFELLIVMIRS 69
QY 81 RRS-----MFLMTC-----KPLYNGPEYI 100
Db 70 RKTGSVTWVNYLASSFLYTKVANAILWAYADRYGLGLFLLCVGMVLPPEPSYRGEHI 129
QY 101 KYF-NDKTIDEELERDKRVTVWVEFFANWSDCQSFAPYADLSLKYNCTGLNFKGVG 159
Db 130 TYFRNAQVFEELARDKTSWLCICFYVWVNSCNVNFAPVFAELSAEYNTDHLKFKGIDIG 189
QY 160 RYTDVGRYKYSTPLTKQLPTLLIFGGKEMRRPQIDKKGRAVSWTFSEENVREFNL 219
Db 190 RFPDVAQKYRISDSFSRQLPTVLIFQOGKEDRRPCVDKSGKLGKFFSSDNVTRATGL 249

QY 220 NELYORA 226
Db 250 NOLYKEA 256

RESULT 5
Q18484 PRELIMINARY; PRT; 265 AA.
AC Q18484;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 30.9 kDa protein.
GN C35D10.10
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2016(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Fulton L.;
RT "The sequence of C. elegans cosmid C35D10.";
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U21324; AAA62566.1;
DR InterPro; IPR000063; ThioRed.
KW Hypothetical protein.
SQ SEQUENCE 265 AA; 30900 MW; 2D0731C464EA5060 CRC64;

Query Match 28.8%; Score 391; DB 5; Length 265;
Best Local Similarity 32.2%; Pred. No. 6e-30;
Matches 84; Conservative 52; Mismatches 81; Indels 44; Gaps 6;

QY 14 VPRLS--RWLAQPYLLSALLSAFLVRLKPLCHGLPTQREDGNPCDFDREVEILMF 71
Db 3 IPRDEVRRALTAFHFFNLLALAPFVIRK-TSLCDYV-FAVEGNEQCEIDSRERILMP 60
QY 72 LSAIVMKNRRS-----MFLMTC-----KPPL 93
Db 61 LLILIAWKGKATNMHYVNNIFLFSKIAGMFLIRADILPGIILYIACLIIVTLVLPPEV 120
QY 94 YMGPIYKIFNDKTTDEELERDKRVTVWVEFFANWSDCQSFAPYADLSLKYNCTGLN 153
Db 121 YNGPEQVTVFOGEQLFEELTRNRNTIWIQFTTWSPECRHTSPVFAELSKQFTLPNMKF 180
QY 154 GKVDYGRYTVSTRYKYSTPLTKQLPTLLIFGGKEMRRPQIDKKGRAVSWTFSEENV 213
Db 181 GKLDIGRWAKEGERFRVNAHPSRQLPTICVFKDAKEIARRLVNDSSRAVPFVFEENC 240
QY 214 IREFNLNELY--QRAKLSKA 232
Db 241 VLAFDLNLYNQEKKGAKA 261

RESULT 6
Q9LYG6 PRELIMINARY; PRT; 253 AA.
ID Q9LYG6
AC Q9LYG6;

```


DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE Hypothetical 28.8 kDa protein.
GN T2P22.30.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL163814; CAB87682.1; -
DR InterPro; IPR000063; ThioRed.
KW Hypothetical protein.
SQ SEQUENCE 253 AA; 28802 MW; 3C09CEA69ACDE42F CRC64;
-Query Match 12.7%; Score 173; DB 10; Length 253;
Best Local Similarity 26.4%; Pred. No. 1.le-08;
Matches 66; Conservative 43; Mismatches 81; Indels 60; Gaps 12;
QY 17 LSRWLAOPYLLSALLSAFLVVK--LPPLCHGLPTQREDGNCDFDWEVE---TLM 70
DB 12 MSRIVDYFLHMFYSYLPKRSSAAPTSHRL-----PD-REIQAFLAFLM 59
QY 71 FLSAIVMKNR-----RSMFLMTCK-----PPLYM---GPEYI 100
DB 60 F-SAIKVVRETEAFVADSLLYAKIELIAYSLIMYRVAVWFSEIVILLAAQPAFS 118
QY 101 KYFNDKID----EELERDKRVT--WIVEFFANNSDCQSFAPYADLSLKYNCTGLNFG 154
DB 119 KLGTAKKLTPQLEDLLSDGNTTYWLYIEFFACSSKVRSSRCFPELSITYNNLSFG 178
QY 155 KVDVGRYTDVSTRYKYSTPLTKQLPTLILFQGGKEAMRRPQ--IDKKGRAVSWTFSEEN 212
DB 179 TVDLGLFPNTAAKFGISLAGMSQLPYILPEKGVESRFFDYVD---APSLPITKKL 235
QY 213 VIREFNLNEL 222
DB 236 LCQHFDLRL 245
RESULT 7
Q8T9N5 PRELIMINARY; PRT; 106 AA.
Q8T9N5
Q8T9N5
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE ThioRedoxin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NMRI;
RA Williams D.L., Heyen J., Alger H.M.;
RT "Molecular and biochemical characterization of thioRedoxin from
RT Schistosoma mansoni";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF473536; AAL79841.1; -
SQ SEQUENCE 106 AA; 11924 MW; CBB0C674ABE2F22F CRC64;
Query Match 10.2%; Score 138.5; DB 5; Length 106;
Best Local Similarity 26.6%; Pred. No. 8e-06;
Matches 32; Conservative 26; Mismatches 26; Indels 13; Gaps 5;
Query Match 10.0%; Score 135.5; DB 2; Length 100;
Best Local Similarity 33.0%; Pred. No. 1.5e-05;
Matches 32; Conservative 26; Mismatches 26; Indels 13; Gaps 5;
QY 100 IKYFNDKT-IDEELERDKRVTWIVEFFANNSDCQSFAPYADLSL-KYNCTGLNFGVD 157
DB 1 MKHITNKAELDQLLTNNKV--VVDYANWCGPKILGPIFEVAQDKDWT---FVKVD 55
QY 158 VGRYTDVSTRYKYSTPLTKQLPTLILFQGGKEAMRR 194
DB 56 VDQANEISSEVEI-----RSIPITFIFFQDCKMADKR 86
RESULT 9
Q9R6P9 PRELIMINARY; PRT; 100 AA.
Q9R6P9
Q9R6P9
AC Q9R6P9;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ThioRedoxin.
GN TRX.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.

Matches 34; Conservative 27; Mismatches 39; Indels 19; Gaps 4;
QY 100 IKYFNDKTIDEELEERDKRVTWIVEFFANNSDCQSFAPYADLSLKYNCTGLNFGKVDVG 159
DB 5 IELKQDGLSELLEQHKNKLVVDVDFATWCGPKTIAPLFKELSEKYDAI---FVKVDVD 61
QY 160 RYTDVSTRYKYSTPLTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSE-ENVIREF 217
DB 62 KLEETARKYNISA-----MPTFAIKNG-----EKVGDVVVGASIAKVEDMIKRF 105
RESULT 8
Q9JN64 PRELIMINARY; PRT; 100 AA.
Q9JN64
Q9JN64
AC Q9JN64;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ThioRedoxin.
GN TRX_L.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A5969;
RX MEDLINE=92038016; PubMed=1718781;
RA Scamrov A., Beabealashvilli R.;
RT "Mycoplasma gallisepticum strain S6 genome contains three regions
RT hybridizing with 16 S rRNA and two regions hybridizing with 23 S and 5
RT S rRNA";
RL FEBS Lett. 291:71-74(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A5969;
RA Skamrov A., Feoktistova E., Goldman M., Beabealashvilli R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; L35043; AAF36768.1; -
DR HSSP; P80579; IQUN.
DR InterPro; IPR000063; ThioRed.
DR Pfam; PF00085; ThioRed. 1.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMS; TIGR01068; thioRedoxin; 1.
DR PROSITE; PS00194; THIOREDOXIN; 1.
KW Redox-active center.
SQ SEQUENCE 100 AA; 11577 MW; C2DFB42A6DBF3C38 CRC64;
Query Match 10.0%; Score 135.5; DB 2; Length 100;
Best Local Similarity 33.0%; Pred. No. 1.5e-05;
Matches 32; Conservative 26; Mismatches 26; Indels 13; Gaps 5;
QY 100 IKYFNDKT-IDEELERDKRVTWIVEFFANNSDCQSFAPYADLSL-KYNCTGLNFGVD 157
DB 1 MKHITNKAELDQLLTNNKV--VVDYANWCGPKILGPIFEVAQDKDWT---FVKVD 55
QY 158 VGRYTDVSTRYKYSTPLTKQLPTLILFQGGKEAMRR 194
DB 56 VDQANEISSEVEI-----RSIPITFIFFQDCKMADKR 86
RESULT 9
Q9R6P9 PRELIMINARY; PRT; 100 AA.
Q9R6P9
Q9R6P9
AC Q9R6P9;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ThioRedoxin.
GN TRX.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.


```
GN Archaeoglobus fulgidus.
OS Archaea: Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Goeyne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AE001015; AAB89961.1; -.
DR HSSP; P00274; 2TRX.
DR TIGR; AF1284; -.
DR InterPro; IPR000063; Thioired.
DR Pfam; PF00085; thioired; 1.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMs; TIGR01068; thioredoxin; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 134 AA; 15347 MW; AB8580DC51D86CC2 CRC64;

Query Match      8.9%; Score 120.5; DB 17; Length 134;
Best Local Similarity 31.4%; Pred. NO. 0.00062;
Matches 27; Conservative 19; Mismatches 31; Indels 9; Gaps 3;

QY 104 NDKTIDELELRKRVTWIWEFFANWSDQSFAPYADLSLKYNTGLNFGKVDVGRYTD 163
Db 36 NSSNFDEFLKKNENV--VVDFAEWCMPCMKIAPVIELAKY-AGKVVFGLKNTDENPT 92
QY 164 VSTRYKVTSPSLTKOLPTLILFQGGK 189
  93 IAARYGISA-----IPTLIFKKGK 112

Search completed: July 9, 2003, 11:57:15
Job time : 83 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: July 12, 2003, 09:52:37 ; Search time 1878 Seconds
(without alignments)
3998.149 Million cell updates/sec

Title: US-09-954-846-2
Perfect score: 1359
Sequence: 1 MAVLAPLIAVYSPRLSRW.....EQPVASTPTVSDGENKRDK 258

oring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+pn.model -DEV=xlh
-O=/cgn2_1/uspto_spoal/us09954846/runat_09072003_102145_12156/app_query.fasta_1.455
-DB=GenEmbl -QWNT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=ext -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09954846.ecgn_1_1687_rnat_09072003_102145_12156 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

tabase : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pt.*
10: gb.to.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rnd.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1326	97.6	1620	6	AX375964 Sequence
2	1326	97.6	1646	9	BC000666 Homo sapi
3	1326	97.6	1669	9	AF132965 Homo sapi
4	1299	95.6	1631	6	AX136115 Sequence
5	1292	95.1	1644	9	AF059753 Homo sapi
6	1220.5	89.8	1978	10	BC013544 Mus muscu
7	1220.5	89.8	200653	2	AC127695 Mus muscu
8	956	70.3	726	6	AX136467 Sequence
9	920	67.7	189271	9	AL138752 Human DNA
10	867	63.8	112389	2	AC073598 Homo sapi
11	579.5	42.6	110116	9	AP001931 Homo sapi
12	579.5	42.6	158349	2	AC009438 Homo sapi
13	579.5	42.6	161397	2	AC090218 Homo sapi
14	579.5	42.6	170540	2	AC092707 Homo sapi
15	579.5	42.6	199992	2	AC021522 Homo sapi
16	522	38.4	130642	2	AP001082 Homo sapi
17	503.5	37.0	149481	2	AC102678 Mus muscu
18	503.5	37.0	216237	2	AC121786 Mus muscu
19	476.5	35.1	1223	3	AY118488 Drosophil
20	476.5	35.1	31857	2	AC020299 Drosophil
21	476.5	35.1	45105	3	AC004311 Drosophil
22	476.5	35.1	160106	3	AC099014 Drosophil
23	476.5	35.1	173347	3	AC099013 Drosophil
24	476.5	35.1	248096	3	AE003795 Drosophil
25	431.5	31.8	5423	6	AX336311 Sequence
26	431.5	31.8	5423	9	AB002382 Human mRN
27	415	30.5	307	6	AX396397 Sequence
28	336	24.7	117711	9	AP000662 Homo sapi
29	336	24.7	153394	2	AP000727 Homo sapi
30	336	24.7	161397	2	AC090218 Homo sapi
31	331.5	24.4	36882	3	U21324 Caenorhabdi
32	330	24.3	153394	2	AP000727 Homo sapi
33	309	22.7	183656	2	AC111385 Rattus no
34	195	14.3	81994	2	AC096003 Rattus no
35	185.5	13.6	81994	2	AC096003 Rattus no
36	185.5	13.6	183656	2	AC111385 Rattus no
37	138.5	10.2	433	3	AF473536 Schistosoma
38	130.5	9.6	1900	6	I01968 Sequence 1
39	130.5	9.6	3552	6	I08049 Sequence 5
40	128.5	9.5	10548	1	AE006828 Sulfolobu
41	127.5	9.4	10733	1	AE013695 Versinia
42	127.5	9.4	220050	1	AJ414156 Versinia
43	126.5	9.3	700	8	AY077707 Fusarium
44	126.5	9.3	110000	2	Continuation (23 o
45	124.5	9.2	1429	3	L28174 Acanthamoeb

ALIGNMENTS

```

/organism="Homo sapiens"
/db_xref="LocusID:51075"
/db_xref="taxon:9606"
/clone="MGC:713 IMAGE:3349733"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_16"

```

	429 a	395 c	376 g	446 t	
BASE COUNT					
ORIGIN					

gnment Scores:		
ed. No.:	2.27e-141	Length: 1646
Score:	1326.00	Matches: 257
Percent Similarity:	87.16%	Conservative: 1
Best Local Similarity:	86.82%	Mismatches: 0
Query Match:	97.57%	Indels: 38
DB:	9	Gaps: 1

US-09-954-846-2 (1-258) x BC000666 (1-1646)

	Qy	1	MetalValLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrp	20
	Db	10	ATGGCGGTCTGGACGACTTAATTCCTCGGTATTCTGGTCCGCAGCTTCACGATGG	69
	Qy	21	LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuValArg	40
	Db	70	CTGCCCAACCTTACTACCTTCTGCGCCTGCTCTGTGCGCTTCTACTCTGTAGG	129

Qy	41	LysLeuProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp	60
Db	130	AAATCTGCCGCGCTGTGCCACGGTCTGCCACCCCAACCGCAAGACGGTAACCCGTTGTGAC	189

Qy	61	PheAspTrpArgGluValuIleuMetPheLeuSerLalaIleValMetMetLysAsn	80
Dd <td>190</td> <td>TTTGACTCGAGAGAAGTGAGATCCTGATGTTTTCAGTCCCATGTGATGATGAAGAAC</td> <td>249</td>	190	TTTGACTCGAGAGAAGTGAGATCCTGATGTTTTCAGTCCCATGTGATGATGAAGAAC	249

Qy	81	ArgArgSer	-----	83
	250	CGAGATCCATC	ACTGTGGAGCAACATATAGGCAACATTTTCATGTTAGTAAAGTGCC	309

Qy	83	-----83
D _b	310	AACACAATCTTTTCTCGCTTGGATATTCGATGGGCTACTTTACATCACA
		CTGCTGC 369

Qy	84	---MetPheLeuMetThrCysLysProLeuTyrMetGlyProGluTyrIleLysTyr	102
Dd	370	AATGTTCCTGATGACGFCGAACCCCCCTATATATGGCCCTGAGTATATCAAGTAC	429

Qy	103	preAsnAspLysThrIleAspGluGluLeuGluArgAspLysArgvalThrTrpIleVal	122
Db	430	TTCAATGATAAAACCATTTGATGAGGAACTAGAACGGGACAAAGAGGCTCACTTGGATTGTG	489

Qy	123	GlupPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu	142
	490	GAGTCTTTGGCCAAATGGTCTAATGAC TGCCAAATCATTTGCCCTATCATGTGACACCTC	549

Qy	143	<p> SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr TCCTCTTAAATACAACTGTACAGGCTAAATTTTCGGAAGCTGGATGTTGGACGCTATACT 550 </p>	162
Db	550	<p> SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr TCCTCTTAAATACAACTGTACAGGCTAAATTTTCGGAAGCTGGATGTTGGACGCTATACT 609 </p>	609

Qy	163	AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu	182
Db	-610	GATGTTATACGGCGGTACAAAGTGAGCATACACCTCACCAAGCAACTCCCTACCCCTG	669

Qy	183	IleLeuPheGlnClyGlyLysGluAlaMetArgProClnIleAspLysLysClyArg	202
Db	670	ATCTGTTCCAAGGTGGCAAGGAGCAATCGCGCGCCACAGATTGACAAAGAACGCG	729

Qy	203	AlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu	222
Db	730	GCTGCTCATGGACCTTCTCTGAGGACAATGTGATCCGACAATTTAACTTTAAATGAGCTA	789
Qy	223	TyrClnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVal	242
Db	790	TACCAGCGGGCCAGAAACATATCAAGGCTGGAGACAAATATCCCTGAGGAGCAGCCTGTG	849
Qy	243	AlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys	258
Db	850	GCTTCAACCCCAACACAGTGTGAGATGGGAAAAACAAGAAGATAAA	897

·RESULT 3
AF132965

AF132965 1669 bp mRNA linear PRI 18-MAY-2000
Homo sapiens CGI-31 protein mRNA, complete cds.

AF132965
AF132965.1 GI:4680700

KEYWORDS
SOURCE
ORGANISM

PREFERENCE

AUTHORS	TITLE
...	...

JOURNAL
MEDLINE

FORMED
REFERENCE
AUTHORS

TITLE
JOURNAL

FEATURES
SOURCE

2
2
2

•

ORIGIN

Pred. No.:
Score:

Percent Sim
Best Local
Query Match

DB:

 \bar{QY}

DB

QY Db

```

QY 41 LysLeuProProLeuCysHisGlyLeuProThrGluArgGluAspGlyAsnProCysAsp 60
Db 170 AACTGCCCGCTCTGCCAGCTGCTCCACCACCAACCGGAGCGGTAAACCGGTGTGAC 229
QY 61 PheAspTrpArgGluValGluLeuLeuMetPheLeuSerAlaIleValMetMetLysAsn 80
Db 230 TTGACTGGAGAGAGTGGAGATCCTGATCTTCTCAGTGCCATTTGTGATGATGAAGAAC 289
QY 81 ArgArgSer----- 83
Db 290 CGCAGATCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTAGTAAAGTGCC 349
QY 83 ----- 83
Db 350 AACACAATTTCTTCCGCTGGATATTCGGATGGCGCTACTTTACATCACACTCTGC 409
QY 84 ---MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102
Db 410 ATAGTCTTCTGATGAGCTGCAAAACCCCTATATATATGATGGCCCTGAGTATATCAAGTAC 469
QY 103 PheAsnAspLysThrIleAspGluLeuLeuGluArgAspLysArgValThrTrpIleVal 122
Db 470 TTCAATGATAAAACCATTTGATGAGGAACCTAGAACGGGACAGAGGGTCACTTGGATTGTG 529
QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
Db 530 GAGTCTTTGCCAATTTGGCTAATGACTGCGCAATCAATTTGCCCTATCTATGCTGACCTC 589
QY 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr 162
Db 590 TCCTTAAATACAACTGACAGGCTAAATTTTGGGAAGGTGATGTTGGACGCTACTACT 649
QY 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
Db 650 GATGTTAGTACGGGTACAAAGTGACACATCACTCCCTCACCAGCAACCTCCCTACCCCTG 709
QY 183 IleLeuPheGlnGlyGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArg 202
Db 710 ATCTCTTCCAGGTGGCAGGAGGCAATCGCGGCGCCACAGATTGACAGAAAGGACGG 769
QY 203 AlaValSerTrpThrPheSerGluGluAsnValIleArgGluPheAsnLeuAsnGluLeu 222
Db 770 GCTGTCTCATGGACCTTCTCTGAGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTA 829
QY 223 TyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProVal 242
Db 830 TACCAGCGGCGCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCGCCCTGTG 889
QY 243 AlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
Db 890 GCTTCAACCCCCACCACAGTGTGATGGGAAACAAAGAGGATAAA 937

JLT 4
.36115
LOCUS 1631 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 37 from Patent EP1067182.
ACCESSION AX136115
VERSION AX136115.1 GI:14272523
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
Hayashi,K.
TITLE Secretory protein or membrane protein
JOURNAL Patent: EP 1067182-A 37 10-JAN-2001;
Helix Research Institute (JP)
FEATURES
source
Location/Qualifiers
1..1631
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```

CDS
16..1134
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC39700.1"
/db_xref="GI:14272524"
LCHLPTQREDGNPCDFDWREILMFSAIVMMKRRSITVEOHIGNIFMFSKVANT
ILFRLDIRMGLLYITLICIVFLMTCKPPLYMGPEYIKYFNDKTIDELEDRKRVTVI
EPFANWSDCQSFAPYIADLSKYNCTGLNFGKVDYGRYTDVSTRYKVSPLTKQLP
TLILFQGGKEAMRRPQIDKKGRAVSWTFSEENYIREFNELNELYORAKKLSKAGNIPE
EQVAOPPPQCMQGRINKILTLAVLLPLSLIPGSPFNHKKPEAAAFYLCFPFGCDWY
GQHAADFKEASRELSCTLOEGLPCCQLFHSKKEIS"
BASE COUNT 411 a 397 c 378 g 445 t
ORIGIN
Alignment Scores:
Pred. No.: 2,71e-138 Length: 1631
Score: 1299.00 Matches: 256
Percent Similarity: 86.82% Conservative: 1
Best Local Similarity: 86.49% Mismatches: 1
Query Match: 95.58% Indels: 39
DB: 6 Gaps: 1
US-09-954-846-2 (1-258) x AX136115 (1-1631)
QY 1 MetaValleLeuAlaProLeuIleAlaLeuValTyrSerValProArgLeuSerArgTrp 20
Db 16 ATGCGGCTCTGGCACCTCTAATTGCTCTGCTGATTGCGTGGCGGCACTTTACGATGG 75
QY 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuLeuValArg 40
Db 76 CTGCCCCAACCTTACTACTTCTGCTCGGCCCTGCTCTGCTGCTTCTTCTACTCTGAGG 135
QY 41 LysLeuProProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysAsp 60
Db 136 AAATGCGCGCGCTCTGCCAGGCTCTGCCACCCCAACGCGAAGCGGTAAACCCGTGTAC 195
QY 61 PheAspTrpArgGluValGluIleLeuMetPheLeuSerAlaIleValMetMetLysAsn 80
Db 196 TTGACTGGAGAGAAGTGGAGATCTCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAC 255
QY 81 ArgArgSer----- 83
Db 256 CGCAGATCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGCC 315
QY 83 ----- 83
Db 316 AACACAATTTCTTCTTCCGCTTGGATATTCGGCATGGCGCTACTTTACATCACACTCTGC 375
QY 84 ---MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyr 102
Db 376 ATAGTGTCTCTGATGAGCTGCAAAACCCCTATATATATGGCCCTGATATATCAAGTAC 435
QY 103 PheAsnAspLysThrIleAspGluLeuGluArgAspLysArgValThrTrpIleVal 122
Db 436 TTCAATGATAAAACCATTTGATGAGGAACCTAGAACGGGACAGAGGGTCACTTGGATTGTG 495
QY 123 GluPhePheAlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeu 142
Db 496 GAGTCTTCTTCCCAATTTGGTCTAATGACTGCTCAATTTGCCCAATTTGCCCTATCTATGCTGACCTC 555
QY 143 SerLeuLysTyrAsnCysThrGlyLeuAsnPheGlyLysValAspValGlyArgTyrThr 162
Db 556 TCCTTTAAATACAACTGATCAGGCGCTAAATTTTGGGAAGGTGATGTTGGACCGCTACTACT 615
QY 163 AspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeu 182
Db 616 GATGTTAGTACGGGTACAAAGTGAGGACATCACTCCCTCACCAGCAACCTCCCTACCTG 675
QY 183 IleLeuPheGlnGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArg 202
Db 676 ATCTGTGTCCAGGTGGCAGGAGGCAATCGCGGCGCCACAGATTTGACAAAGAGGACGG 735

```



```

...
Db 59008 AAGTAATGCGGAGCAGACAGACTGACAAAGAGCGGCTGTCTCATGACCTTCTCT 58949
QY 210 AGUGluasnValIleArgGluPheAsnLeuAsnGluLeuTyrGlnArGAlaLysLysLeu 229
Db 58948 GAGGAGAATGTGATCCGAGAATTAACTTAATGAGCTATACACGCGGGCAAGAGCTA 58889
QY 230 SerLysAlaGlyAspAsnIleProGluGlnProValAlaSerThrProThrVal 249
Db 58888 TCAAGGCTGTGAGACATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCAACACAGTG 58829
QY 250 SerAspGlyGluAsnLysLysAspLys 258
Db 58828 TCAGATGGGGAAGACAGAGAGGATAAA 58802

RESULT 10
AC073598 Homo sapiens chromosome 10 clone CTC-268N23, WORKING DRAFT
LOCUS AC073598
DEFINITION AC073598.3 GI:20279338
ACCESSION AC073598
VERSION AC073598.3
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
~URCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 112389)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroli,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B.,
Homs1,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,X., Johnson,R., Jolivet,S., Joudah,S.,
Karissoun,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokkwo,S., Oguh,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubenok,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,K.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
Direct Submission
Unpublished
REFERENCE 2 (bases 1 to 112389)
Worley,K.C.
TITLE
Direct Submission

```

```

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BASE COUNT 29689 a 25997 c 25893 g 29967 t 843 others
ORIGIN
Alignment Scores:
Pred. No.: 1.17e-86 Length: 112389
Score: 867.00 Matches: 181
Percent Similarity: 74.70% Conservative: 5
Best Local Similarity: 72.69% Mismatches: 19
Query Match: 63.80% Indels: 45
DB: 2 Gaps: 3
US-09-954-846-2 (1-258) x AC073598 (1-112389)

```

```

Submitted (26-JUN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 112389)
Worley,K.C.
Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Apr 24, 2002 this sequence version replaced gi:16118028.
-----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: HMRT
Center clone name: CTC-268N23
-----
Summary Statistics
Sequencing vector: M13;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 10470 bases at least Q40
Consensus quality: 107539 bases at least Q30
Consensus quality: 109442 bases at least Q20
Estimated insert size: 111011; sum-of-ctigs estimation
Quality coverage: 3x in Q20 bases; sum-of-ctigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2191: contig of 2191 bp in length
* 2192 2291: gap of unknown length
* 2292 5898: contig of 3607 bp in length
* 5899 5998: gap of unknown length
* 11406: contig of 5408 bp in length
* 11506: gap of unknown length
* 11507 21739: contig of 10233 bp in length
* 21740 21839: gap of unknown length
* 21840 31956: contig of 10117 bp in length
* 31957 32056: gap of unknown length
* 32057 46774: contig of 14718 bp in length
* 46775 46875: gap of unknown length
* 46875 68219: contig of 21344 bp in length
* 68219 68318: gap of unknown length
* 68319 85954: contig of 17636 bp in length
* 85955 86054: gap of unknown length
* 86055 112389: contig of 26335 bp in length.
FEATURES
Location/Qualifiers
1..112389
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="CTC-268N23"
BASE COUNT 29689 a 25997 c 25893 g 29967 t 843 others
ORIGIN
Alignment Scores:
Pred. No.: 1.17e-86 Length: 112389
Score: 867.00 Matches: 181
Percent Similarity: 74.70% Conservative: 5
Best Local Similarity: 72.69% Mismatches: 19
Query Match: 63.80% Indels: 45
DB: 2 Gaps: 3
US-09-954-846-2 (1-258) x AC073598 (1-112389)

```

```
QY 53 ArgGluAspGlyAsnPro-----CysAspPheAspTrpArgGluValGluIleLeu 69
Db 19522 AGAAGAGACTCGGAACCCCAAGAGAGTTCGCCGAACCGGAGAGAGCGGAGATCCTG 19581
QY 70 MetPheLeuSerAlaIleValMetMetLysAsnArgSer-----83
Db 19582 ATGTTTCTCAGTGCCATTGTGATGATGAAGAACCAGCATCCACTACTGTGGAGCAACAT 19641
QY 83 -----83
Db 19642 ATAGGCAACATTTCATGTTTAGTAAAGTGGCAACGCAATTCCTTTCTCCGCTTGGAT 19701
QY 84 -----MetPheLeuMetThrCysLys---90
19702 ATTCCGATGGGCTACTTTACATCACACTCTGCATAGTGTCTCTGATGAGCGTCAACACC 19761
91 ---ProProLeuTyrMetGlyProGluTyrIleLysTyrPheAsnAspLysThrIleAsp 109
Db 19762 CCCCTCCCCA-TATATGGCCCTGAGTATATCAAGTACTTCAATGATAAACCATTGAT 19820
QY 110 GluGluLeuGluArgAspLysArgValThrTrpIleValGluPheAlaAsnTrpSer 129
Db 19821 GAGGAACCTAGAACGGGACACACAGGTCACCTGGATTGTGGAGTCTTTGCCAATTGGTCT 19880
QY 130 AsnAspCysGlnSerPheAlaProIleTyrAlaAspLeuSerLeuLysTyrAsnCysThr 149
Db 19881 AATGACTGCCTCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAAGTACAACCTGACA 19940
QY 150 GlyLeuAsnPheGlyLysValAspValGlyArgTyrThrAspValSerThrArgTyrLys 169
Db 19941 GGGCTAAATTTGGGACGGTGGATGCTGGACGCTACTAGTATGATGCTGACCGGTACAAA 20000
QY 170 ValSerThrSerProLeuThrLysGlnLeuProThrLeuIleLeuPheGlnGlyLys 189
Db 20001 GTGAGCACATCACCCCTCACCAATCAACTCCCTACCTGATCTCTGTCCAAAGGTGCAAG 20060
QY 190 GluAlaMetArgArgProGlnIleAspLysLysGlyArgAlaValSerTrpThr-Phese 209
Db 20061 AAGGTAATCGCGGACGACAGACTGANCAGAAGGACGGGCTGCTCATGGAGCTTTCTC 20120
QY 209 rGluAsnValIleArgGluPheAsnLeuAsnGluLeuTyrGlnArgAlaLysLysLe 229
Db 20121 TGAGGAGATGATCCGAGATTTAACTTAATGAGCTATACACGCGGGCAGAGCT 20180
QY 229 uSerLysAlaGlyAspAsnIleProGluGlnProValAlaSerThrProThrTrpVa 249
Db 20181 TTCAAAGGCTGGAGACAATATCCCTGAGGAGCAGGCTGGGCTTTCAGCCCCACACAGT 20240
QY 249 lSerAspGlyGluAsnLysLysAsp 257
Db 20241 GTCAGATGGGAAAGCAAGAGGAT 20265
RESULT 11
AP001931
LOCUS AP001931 110116 bp DNA linear PRI 28-MAY-2002
DEFINITION Homo sapiens genomic DNA, chromosome 11q clone:RP11-691N7, complete
sequences.
ACCESSION AP001931
VERSION AP001931.5 GI:21218128
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
ToTokI,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2000)
REFERENCE 2 (bases 1 to 110116)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., ToTokI,Y., Watanabe,H. and Sakaki,Y.
```

```
Direct Submission
Submitted (28-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gs.c.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On May 27, 2002 this sequence version replaced gi:20334325.
FEATURES
Location/Qualifiers
source
1..110116
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-691N7"
BASE COUNT 29082 a 23179 c 24671 g 33184 t
ORIGIN
Alignment Scores:
Pred. No.: 7 5e-54 Length: 110116
Score: 579.50 Matches: 172
Percent Similarity: 27.37% Conservative: 1
Best Local Similarity: 27.22% Mismatches: 2
Query Match: 42.64% Indels: 459
DB: 9 Gaps: 4
US-09-954-846-2 (1-258) x AP001931 (1-110116)
QY 84 MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysPhe 103
Db 18410 GTCTTCTGATGACGTGCAACCCCTATATATGGCCCTGAGTATATCAAGTACTTC 18469
QY 104 AsnAspLysThrIleAspGlu-----110
Db 18470 AATGATAAAACCATTCATGT-GAGTGCTCTTTCCCTTTCTGTTCTTGGTGCTCTTGTG 18528
QY 110 -----110
Db 18529 GGTGATTTTGTAGTTGTGCTCTCCATCTAGGAGGAACAACAGCTCTCAAAATGAT 18588
QY 110 -----110
Db 18589 GTCACGGGCACGTGTGTTTATTATGAGCGCGGGGAGAGGACACACGAGTACGAGNAT 18648
QY 110 -----110
Db 18649 TAGATGCTAAAGTCTGAACCTCCAGGAGGCTATGTGGATCCAGCTGACTTTTCTTCCC 18708
QY 111 -----GluLeuGluArgAspLysArgValThrTrpIleValGluPhePhe 125
Db 18709 TGTATTGGCAGGAGGAACCTAGAACGGGACAGAGGCTCATTGGATTGTGGAGTCTTT 18768
QY 126 AlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeuSerLeuLys 145
Db 18769 GCCAATTGGTCTAATGACTGCCAATCATTTGCCCTTATCTATGCTGACCTCTCCCTTAA- 18827
QY 145 -----145
Db 18828 GTGAGTAGTGCAAAAGGAGGAGTGTGGAATGGAGATGCTGTGCCTTCCCTCTCACTGT 18887
QY 145 -----145
Db 18888 TTTTGGCTTTTCTTTTCTTTTCTTTTGGCCTTGATTTTTCACACATGGTAACCAAGCATCTCC 18947
QY 145 -----145
Db 18948 CTCCTCCCTCTTAAATATCTATACTTCCACTTTCTTGTGATCCATTTATTTTTCAGCAT 19007
QY 146 -----TyrAsnCysThrGlyLeuAsnPheGlyLysValAsp 157
Db 19008 ATTTAAATAATATATTTCTTTTTCAGATACAACTGTACAGGGCTAAATTTTGGGAAGGTGAT 19067
QY 158 ValGlyArgTyrThrAspValSerThr-----166
|||||
```

Db 19068 GTTGGAGCTTACTATGATGTTAGTACGCGGTATGTAAGACCTGGCGAGAGGGTCTGAGC 19127
QY 167 -----Ar 167
Db 19128 AGGGAATACATTTGAGTGATACATACAGGAGACATTTAGAGAACTTTCTGGGCCCTGCAG 19187
QY 167 qYrLYsValSerThrSerProLeuThrLYsGlnLeuProThrLeuLeuPheGlnG1 187
Db 19188 GTACAAGTGAGCAGATCCCTCCACCAAGCACTCCCTACCTGATCTCTGTTCACAGG 19247
QY 187 yGLyLYsGluAlaMetArgArgProGlnIleAspLYsLYsGlyArgAlaValSerTrpTh 207
Db 19248 TGGCAAGGAGGCAATCGCGGGCCACAGATTTGACAAAGGAGCGGCTGTCTCATGGAC 19307
QY 207 rPheSerGlu-----210
Db 19308 CTTCTCTGAGGTACCTGAAGGAAGGCGCAGGTGCATGAAGGGTGAGAACAGTAGGTGGG 19367
QY 210 -----210
Db 19368 CTTTCAAGCCCTACCCGGGTTTGATTCACAGCTCTGCCACTTGCCCAATTAGCTTTGAGGG 19427
QY 210 -----210
Db 19428 TGTGGACTAGTTACTAGACCTCATGTTTAAATTCATTTAGCAAAATATATATTATGTCAG 19487
QY 210 -----210
Db 19488 ACCTGCTAGTACTGGGTAAACAATCATTAGTTAAAGATTGGTCTCAGCGCCAGCGCTGG 19547
QY 210 -----210
Db 19548 CTCACCCCTGTAATCCAGCACATTTGGGAGCGCGAGCTGGCAGATCACCTGAGGTGAGG 19607
QY 210 -----210
Db 19608 AGTTTGAGACCGCTGACCAACATGAGAAACCCCGTGTCTACTAAATAACAATA 19667
QY 210 -----210
Db 19668 GCCGAGTGTGGTGGCGCATGCTATATATGCCAGTATTTGGGAGGCTGAGGAGGAGAA 19727
QY 210 -----210
Db 19728 CGTTGAACCCGGAGGTGGAGTTGCGAGTGAGATTCACCAATTCGGGTCCAGCC 19787
QY 210 -----210
Db 19788 TGGGCAACAAGAGACTCCATCAGGAAAAAAGAAAGATTCTGTCTCGAAGGCCACTGC 19847
QY 210 -----210
Db 19848 ACCTATTGCTGTAGGTGTAGGTTCCTGCTCAGTTAAATTGGAGGTAATAACCCCTGGCT 19907
QY 210 -----210
Db 19908 TACAGGGTGTGTTGCGAGATTACATTAATAATTAGGTATATCTTGGCAGCGTGATT 19967
QY 210 -----210
Db 19968 GGCACATAATGGCTATTTAATTAATTTAGTTGACTTTTCTGGTACCTAAAGAGAGAGG 20027
QY 210 -----210
Db 20028 TTAGGAAGATTTTGTCTCTTCTTACTCTCTCTTCCAGACTTTGTGTAAATACC 20087
QY 210 -----210
Db 20088 TCTTACTTCCAGGCTTTTACTCTCCCTTCCAAACCCAGATCTCTGAGTGTGCATCTCT 20147
QY 211 -----GluAsnValIleArgGluPheAsnLeuAsnGluLeuThrGlnArgAlaLYsL 228
Db 20148 TTGTGCGAGAGAAATGTGATCCGAGAAATTTAACTTAATGAGCTATACCGGGGCCAAGA 20207

QY 228 ySLeuSerLYsAlaGlyAspAsnIleProGluGluGlnProValAlaSerThrProThr 248
Db 20208 AACATATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCTGTGCTTCAACCCCA 20267
QY 248 hrValSerAspGlyGluAsnLYsLYsAspLYs 258
Db 20268 CAGTGTCAAGTGGGAAACACAGAGGATAAA 20299

AC009438 158349 bp DNA linear HTG 04-SEP-2000
AC009438 Homo sapiens chromosome 11 clone RP11-77M17 map 11, WORKING DRAFT
DEFINITION
LOCUS
SEQUENCE, 25 unordered pieces.

AC009438 3 GI:9966254
VERSION
KEYWORDS
SOURCE
ORGANISM
HTG: HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens.
Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
1 (bases 1 to 158349)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone RP11-77M17
Unpublished
2 (bases 1 to 158349)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
Funke,R., Gage,B., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,B., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrum,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.

Direct Submission
Submitted (22-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 4, 2000 this sequence version replaced gi:765347.
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WBIR

Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L1359
Center clone name: 77_M17

----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-primer-amerham; 5% of reads
Chemistry: Dye-terminator Big Dye; 95% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 138242 bases at least Q40
Consensus quality: 145728 bases at least Q30
Consensus quality: 149650 bases at least Q20
Insert size: 172000; agarose-fp
Insert size: 155949; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; agarose-fp
Quality coverage: 4.4 in Q20 base.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1 3650: contig of 3650 bp in length
* 3651 3750: gap of 100 bp
* 3751 4867: contig of 1117 bp in length
* 4868 4967: gap of 100 bp
* 4968 6757: contig of 1790 bp in length
* 6758 6857: gap of 100 bp
* 6858 7991: contig of 1134 bp in length
* 7992 8091: gap of 100 bp
* 8092 9216: contig of 1125 bp in length
* 9217 9316: gap of 100 bp
* 9317 10384: contig of 1068 bp in length
* 10385 10484: gap of 100 bp
* 10485 11835: contig of 1351 bp in length
* 11836 11935: gap of 100 bp
* 11936 13429: contig of 1494 bp in length
* 13430 13529: gap of 100 bp
* 13530 15092: contig of 1563 bp in length
* 15093 15192: gap of 100 bp
* 15193 17380: contig of 2188 bp in length
* 17381 17480: gap of 100 bp
* 17481 20586: contig of 3106 bp in length
* 20587 20686: gap of 100 bp
* 20687 24958: contig of 4272 bp in length
* 24959 25058: gap of 100 bp
* 25059 30271: contig of 5213 bp in length
* 30272 30371: gap of 100 bp
* 30372 35676: contig of 5305 bp in length
* 35677 35776: gap of 100 bp
* 35777 40640: contig of 4864 bp in length
* 40641 40740: gap of 100 bp
* 40741 46368: contig of 5628 bp in length
* 46369 46468: gap of 100 bp
* 46469 53373: contig of 6905 bp in length
* 53374 53473: gap of 100 bp
* 53474 60571: contig of 7098 bp in length
* 60572 60671: gap of 100 bp
* 60672 67052: contig of 6381 bp in length
* 67053 67152: gap of 100 bp
* 67153 75831: contig of 8679 bp in length
* 75832 75931: gap of 100 bp
* 75932 84118: contig of 8187 bp in length
* 84119 84218: gap of 100 bp
* 84219 95075: contig of 10857 bp in length
* 95076 95175: gap of 100 bp
* 95176 110669: contig of 15494 bp in length
* 110670 110769: gap of 100 bp
* 110770 139814: contig of 29045 bp in length
* 139815 139914: gap of 100 bp
* 139915 158349: contig of 18435 bp in length.
  
```

FEATURES

source

```

1..158349
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /map="11"
  /clone="RP11-77M17"
  /clone_lib="RPC1-11 Human Male BAC"
misc_feature
  1..3650
  /note="assembly_fragment
  clone_end:SP6
  vector_side:left"
misc_feature
  3751..4867
  /note="assembly_fragment"
misc_feature
  4968..6757
  /note="assembly_fragment"
misc_feature
  6858..7991
  /note="assembly_fragment"
misc_feature
  8092..9216
  /note="assembly_fragment"
misc_feature
  9317..10384
  /note="assembly_fragment"
  
```

```

misc_feature
  /note="assembly_fragment"
  10485..11835
  /note="assembly_fragment"
  11936..13429
  /note="assembly_fragment"
  13530..15092
  /note="assembly_fragment"
  15193..17380
  /note="assembly_fragment"
  17481..20586
  /note="assembly_fragment"
  20687..24958
  /note="assembly_fragment"
  25059..30271
  /note="assembly_fragment"
  30372..35676
  /note="assembly_fragment"
  35777..40640
  /note="assembly_fragment"
  40741..46368
  /note="assembly_fragment"
  46469..53373
  /note="assembly_fragment"
  53474..60571
  /note="assembly_fragment"
  60672..67052
  /note="assembly_fragment"
  67153..75831
  /note="assembly_fragment"
  75932..84118
  /note="assembly_fragment"
  84219..95075
  /note="assembly_fragment"
  95176..110669
  /note="assembly_fragment"
  110770..139814
  /note="assembly_fragment"
  139915..158349
  /note="assembly_fragment
  clone_end:T7
  vector_side:right"
  
```

BASE COUNT 40315 a 36712 c 36032 g 42881 t 2409 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.19e-53 Length: 158349
 Score: 579.50 Matches: 172
 Percent Similarity: 27.37% Conservative: 1
 Best Local Similarity: 27.22% Mismatches: 2
 Query Match: 42.64% Indels: 459
 DB: 2 Gaps: 4

US-09-954-846-2 (1-258) x AC009438 (1-158349)

```

QY      84 MetPheLeuMetThrCysLysProProLeuTyMetClyProGluTyrIleLysTy:Phe 103
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      143071 GTGTTCTTGATGACGTGCAAAACCCCTATATATGGGCCCTGAGTATATCAAGTACTTC 143130
       |||
QY      104 AsnAspLysThrIleAspGlu-----
       |||
Db      143131 AATGATAAAACCATTCATGTG-TGACTGCTCTTTCCCTCTTCTGTTGTCCTCTGTG 143189
       |||
QY      110 -----
       |||
Db      143190 GGTGATTTTGACTGTGCTCTCCATTCAGGAGGAACAACAGTGCTTCAAAATGGAT 143249
       |||
QY      110 -----
       |||
Db      143250 GTCACGGGCACTGTGTTTCATTATGAGACCGGGGAGAGGAACACAGGATGAGGAAT 143309
       |||
QY      110 -----
       |||
Db      143310 TAGATGCTAAAGTCTGAACCTTCCCAGGAGGCTATGTGGATCCAGCTGACTTTCTTCCC 143369
       |||
  
```

```
QY 111 -----GluLeuGluArgAspLysArgValThrIleValGluPhePhe 125
Db 143370 TGTATTTGGCAGGAGGAGTAAAGCGGACAGAGGCTCACTTGGATTGTGGAGTCTTT 143429
QY 126 AlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeuSerLeuLys 145
Db 143430 GCCAATTTGGTCTAATGACTGCGCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAA- 143488
QY 145 ----- 145
Db 143489 GTGAGTAGTGCAAAGGGGAGGATGGTGAATGGAGATGCTGTGCCCTCCCTCTCACGTGT 143548
QY 145 ----- 145
Db 143549 TTTTGGCTTTTCTTTTCTTTTGGCCTTGATTTTTCACACATGGTAACAAAGGCATCTCC 143608
QY 145 ----- 145
Db 143609 CTCTCCCTCTTAAATATCTATCTTCCACTTCTTCCCTTGATCCATTAATTTTTTTCAGCAT 143668
QY 146 -----TyrAsnCysThrGlyLeuAsnPheGlyLysValAsp 157
Db 143669 ATTAATAATATATTTCTTTTCAGATACAACTGTACAGGCTAAATTTTGGGAAGTGGAT 143728
QY 158 ValGlyArgTyrThrAspValSerThr----- 166
Db 143729 GTTGGACGCTATCTGATGTGTAGTACGCGGTATGTAAGACCTGGCAGAGGGTCTGAGC 143788
QY 167 -----Ar 167
Db 143789 AGGAAATCACTTTGAGTGATACATACAGGGACATTTAGAGAACTTTCTGGGCCCTGCGAG 143848
QY 167 qTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeuIleLeuPheGlnG1 187
Db 143849 GTCAAAGTGAGCATACACCCCTCACCAACCACTCCCTACCTGATCTCTGTCTCCAGG 143908
QY 187 yGlyLysGluAlaMetArgArgProGlnIleAspLysGlyArgAlaValSerTrpTh 207
Db 143909 TGGCAAGGAGGCAATGGCGGCCACAGATTGACAAGAAGAGCGGGTGTCTCATGGAC 143968
QY 207 rPheSerGlu----- 210
Db 143969 CTCTCTGAGGTACTGAAAGGAGGGCAGGTGCATGAGGGTGACAGACAGTAGTGGG 144028
QY 210 ----- 210
Db 144029 CTTTCAAGCCCTACCCGGTTTGATTACAGCTCTGCCACTTGCCCATTTAGCTTTGAGGG 144088
QY 210 ----- 210
Db 144089 TGTGGACTAGTTACTAGACCTCATGTTTAAATTCATTAGCAAAATATATATTATGTGCAG 144148
QY 210 ----- 210
Db 144149 ACCTGCTAGTACTGGGTAAACATCATTTAGTTAAGATTTGGTCTCAGGCCAGCAGGTGG 144208
QY 210 ----- 210
Db 144209 CTCACGCTGTATATCCAGCATTTTGGAGGCCGAGGGTGCAGATCACCTGAGGTGAGG 144268
QY 210 ----- 210
Db 144269 AGTTTGAACAGCCTGACCAACATGGAGAAACCCCGTGTCTACTAAATAATACAAAATA 144328
QY 210 ----- 210
Db 144329 GCCGAGTGTGGTGGCGCATGCTTATAATGCCAGCTATTTGGGAGGCTGAGCGAGGAGAT 144388
QY 210 ----- 210
Db 144389 CGCTTGAACCGGGAGGTGGAGGTTGCGTGAAGTTCAGCTGAGATTGCACCATTTGGTCCAGCC 144448

QY 210 ----- 210
Db 144449 TGGCAACAAGAGACTCCATCAGGAAAAAAGAGATTCTGTCTGAAGGCCACTGC 144508
QY 210 ----- 210
Db 144509 ACCTATTGCTGTAGGTCTTAGGTCCCTGTCTAGTTAATTTGGAGGTAAATACCCCTGGCT 144568
QY 210 ----- 210
Db 144569 TACAGGTTGTTTTCAGATTACATTAATAATTTAGGTATAAATTTCTTGGCAGCGTGAAT 144628
QY 210 ----- 210
Db 144629 GGCACATAATGGGTATTTAATTAATTTAGTTGACTTTTCTGGGTACCTAAAAAGAGAGG 144688
QY 210 ----- 210
Db 144689 TTAGGAAGATTTTGTCTTGTCTTACTCTCTTCTCCAGACTTTTGTGTAAATACC 144748
QY 210 ----- 210
Db 144749 TCTTACTTCCAGGCTCTTTACTCTCTCCCTTCCACCCAGATCCTGACGTGTGCATCTCT 144808
QY 211 -----GluAsnValIleArgGluPheAsnLeuAsnGluLeuTyrGlnArgAlaLysL 228
Db 144809 TTGTGCAGGAATGTGTATCCGAGAATTTAACTTAAATGAGCTATACAGCGGGCCAAGA 144868
QY 228 ysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProValAlaSerThrProThr 248
Db 144869 AACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCAACCA 144928
QY 248 hrValSerAspGlyGluAsnLysLysAspLys 258
Db 144929 CAGTGTCAAGTGGGGAACCAAGAGGATAAA 144960

RESULT 13
AC090218/c
LOCUS
DEFINITION
AC090218.3 GI:14718351
AC090218
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

AC090218 161397 bp DNA linear HTG 14-JUL-2001
Homo sapiens chromosome 11 clone RP11-691N7 map 11, WORKING DRAFT
SEQUENCE, 17 unordered pieces.
AC090218
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Homo sapiens
Homo sapiens
Mammalia; Eukarya; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eukarya; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 161397)
Homo sapiens chromosome 11, clone RP11-691N7
2 (bases 1 to 161397)
Unpublished
Biren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Boguslavskiy, L., Bouckhgalter, B., Brown, A.,
Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T.,
Lehocky, J., Levine, R., Liu, G., MacLean, C., Macdonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
Sougniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,
```

Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 14, 2001 this sequence version replaced gi:13940674:
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L12248
 Center clone name: 691_N_7
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 155030 bases at least Q40
 Consensus quality: 157856 bases at least Q30
 Consensus quality: 159050 bases at least Q20
 Insert size: 157000; agarose-fp
 Insert size: 159797; sum-of-contigs
 Quality coverage: 12.0 in Q20 bases; agarose-fp
 Quality coverage: 11.8 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 17 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 7057: contig of 7057 bp in length
 * 7058 7157: gap of 100 bp
 * 7158 7519: contig of 362 bp in length
 * 7520 7619: gap of 100 bp
 * 7620 8839: contig of 1220 bp in length
 * 8840 8939: gap of 100 bp
 * 8940 9986: contig of 1047 bp in length
 * 9987 10086: gap of 100 bp
 * 10087 11612: contig of 1526 bp in length
 * 11613 11712: gap of 100 bp
 * 11713 13287: contig of 1575 bp in length
 * 13288 13387: gap of 100 bp
 * 13388 14435: contig of 1048 bp in length
 * 14436 14535: gap of 100 bp
 * 14536 16619: contig of 2084 bp in length
 * 16620 16719: gap of 100 bp
 * 16720 17857: contig of 1138 bp in length
 * 17858 17957: gap of 100 bp
 * 17958 19972: contig of 2015 bp in length
 * 19973 20072: gap of 100 bp
 * 20073 22132: contig of 2060 bp in length
 * 22133 22323: gap of 100 bp
 * 22323 58899: contig of 36667 bp in length
 * 58900 58999: gap of 100 bp
 * 59000 70617: contig of 11618 bp in length
 * 70618 70717: gap of 100 bp
 * 70718 89521: contig of 18804 bp in length
 * 89522 89621: gap of 100 bp
 * 89622 118253: contig of 28632 bp in length
 * 118254 118353: gap of 100 bp
 * 118354 155823: contig of 37470 bp in length
 * 155824 155923: gap of 100 bp
 * 155924 161397: contig of 5474 bp in length.
 * Location/Qualifiers
 1.161397
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

FEATURES
source

/chromosome="11"
 /map="11"
 /clone="RP11-691N7"
 /clone_lib="RPC1-11 Human Male BAC"
 1..7057
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"
 7158..7519
 /note="assembly_fragment"
 7620..8839
 /note="assembly_fragment"
 8940..9986
 /note="assembly_fragment"
 10087..11612
 /note="assembly_fragment"
 11713..13287
 /note="assembly_fragment"
 13388..14435
 /note="assembly_fragment"
 14536..16619
 /note="assembly_fragment"
 16720..17857
 /note="assembly_fragment"
 17958..19972
 /note="assembly_fragment"
 20073..22132
 /note="assembly_fragment"
 22233..58899
 /note="assembly_fragment"
 59000..70617
 /note="assembly_fragment"
 70718..89521
 /note="assembly_fragment"
 89622..118253
 /note="assembly_fragment"
 118354..155823
 /note="assembly_fragment"
 155924..161397
 /note="assembly_fragment
 clone_end:T7
 vector_side:right"
 BASE COUNT 43394 a 34530 c 35670 g 46202 t 1601 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.21e-53 Length: 161397
 Score: 579.50 Matches: 172
 Percent Similarity: 27.37% Conservative: 1
 Best Local Similarity: 27.22% Mismatches: 2
 Query Match: 42.64% Indels: 459
 DB: 2 Gaps: 4
 US-09-954-846-2 (1-258) x AC090218 (1-161397)
 QY 84 MetPheLeuMetThrCysLysProLeuTyMetGlyProGluTyrlleLysTyPhe 103
 Db 38287 GTGTTCTCGATGACGTGCAACCCCTATATATGGGCCCTGAGTATATCAAGTACTTC 38228
 QY 104 AsnAspLysThrIleAspGlu----- 110
 Db 38227 AATGATAAAACCATGATGT-GAGTGTCTCTTTCCCTTCTGTTCCTGGTCCCTTGTG 38169
 QY 110 ----- 110
 Db 38168 GGTGATTTTGTAGTTGTCTCTCCATTCCTAGGAGGAACAACAGTGCTTCAAAATGGAT 38109
 QY 110 ----- 110
 Db 38108 GTCACGGGCACCTGTGTTTCATTATGAGAGCCGGGAGAGGAACACAGGATGAGGAAT 38049
 QY 110 ----- 110

Db	38048	TAGATGCTAAAGCTCTGAACCTTCCAGGAGGCTATGTGGATCCAGCTGACCTTTTCTTCCC	37989
Qy	111	-----GluLeuGluArgAspLysArgValThrTrpIleValcIuPhePhe	125
Db	37988	TCTATTTGGCAGGAGGAACCTAGACGGGACAGAGGGTCACCTTGGATTGTGGAGTTCCTT	37929
Qy	126	AlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyYrAlaAspLeuSerLeuLys	145
Db	37928	GCACATTTGGTCTAAATGACTGCGCAATCATTTGCCCTATCTATGTGTGACCTCTCCCTTAA-	37870
Qy	145	-----	145
Db	37869	GTGAGTAGTGCAAAAGGAGGGATGTGGNAATGGAGATGCTGTGCCTTCCCTCTCCTACTGT	37810
Qy	145	-----	145
Db	37809	TTTTGGCTTTTCTTTTCTTTTGGCCTTGATTTTTCACACATGTTAAACCAAGGCATCTCC	37750
Qy	145	-----	145
Db	37749	CTCTCCCTCTTAAATATCTATACTTCCACTTTCCTTGATCCCATTTATTTTTCAGCAT	37690
Qy	146	-----TyrAsnCysThrGlyLeuAsnPhesGlyLysValAsp	157
Db	37689	ATTAAATAATATATCTTTTCAGATACAACTGTACAGGGCTAAATTTTGGGAAGGTGAT	37630
Qy	158	ValGlyArgTyThrAspValSerThr-----	166
Db	37629	GTGGACGCTTACTGATGTTTAGCGCGGTATGTAAAGACCTGGGCAGAGGGTCTGAGC	37570
Qy	167	-----Ar	167
Db	37569	AGGAAATCACTTTGAGTGATACATACAGGGACATTTAGAGAACTTTCTGGGCCCTGCAG	37510
Qy	167	gTyLysValSerThrSerProLeuThrLysGlnLeuProThrLeuIleLeuPheGlnG1	187
Db	37509	GTACAAAGTGAGCAGATCACCCCTCACCAAGCACTCCCTACCTGATCTCTGTTCCTCAAGG	37450
Qy	187	yGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArgAlaValSerTrpTh	207
Db	37449	TGCAAGGAGGCAATGCGCGCGCCACAGATTGCACAAAGACGGGCTGTCTCATGGAC	37390
Qy	207	rPheSerGlu-----	210
Db	37389	CTTCTCTGAGGTACCTGAAAGGAGGCGAGGTGCATGAAGGTGCAGAACAGTAGGTGGG	37330
Qy	210	-----	210
Db	37329	CTTCAAGCCCTACCGGGTTTGTATCACAGCTCTGCCACTTGCCCATTAGCTTGTAGGG	37270
Qy	210	-----	210
Db	37269	TGTGGACTAGTTACTAGACCTCATGTTTAAATTCATTTAGCAAAATATATATATGTGCAG	37210
Qy	210	-----	210
Db	37209	ACCTGTAGTACTGGGTAACAATCATTTAGTTAAGATTGGTCTCAGGCCAGCAGCGTGG	37150
Qy	210	-----	210
Db	37149	CTCACGCTGTAAATCCACGACTTTTGGAGCGGAGGGCTGGCAGATCACCTGAGGTGAGG	37090
Qy	210	-----	210
Db	37089	AGTTTGAGACGCGCTGACCAACATGGAGAACCCCGCTGTCTACTAAAAATACAAAATA	37030
Qy	210	-----	210
Db	37029	GCCGAGTGTGGTGGCCATGCCCTAATATGCCAGCTATTGGGAGGCTGAGGCAGGAGAT	36970
Qy	210	-----	210
Db	36969	CGCTTGAACCCGGAGGTGGAGGTTCAGTGAGCTGAGATTGCACCAATTGGCGTCCAGCC	36910

Qy	210	210
Qy	210	210
Db	36909	TGGGCAACAAGAGACTCCATCAGGAAAAAAGAGATTCGTCTCGAAGGCCACTGC 36850
Qy	210	210
Db	36849	ACCTATTGCTGTAGGTGTAGGTTCCCTGTCAGTTAATTGGAGGTAATAACCCCTGGCT 36790
Qy	210	210
Db	36789	TACAGGGTGTTCAGAGATTACAATAAATTTAGGTATAATTTCTGGCACGGTGATT 36730
Qy	210	210
Db	36729	GGCACAATAGGGTATTAACTATTAGTTGACTTTTCTGGTGACCTAAAAAGAGGAAG 36670
Qy	210	210
Db	36669	TTAGGGAAGATTTTGTCTCTTGTCTTACTCTCTTCCCTTCCACCCAGATCTTCCAGACATTGTGTAATAATACC 36610
Qy	210	210
Db	36609	TCTTACTTCCAGGCTCTTACTCTCTCTTCCCTTCCACCCAGATCTGACGTGTCATCTCTT 36550
Qy	211	-----GluAsnValIleArgGluPheAsnLeuAsnGluLeuTyrglnArgAlaLysL 228
Db	36549	TTGTGACGAGGAATGTGATCCGAGATTTTAACTTAATGAGCTATACCGCGGCCAAGA 36490
Qy	228	ysLeuSerLysAlaGlyAspAsnIleProGluGlnProValAlaSerThrProHrt 248
Db	36489	AACATCAAGAGCTGGAGACATATCCCTTGAGGAGCAGCTGTGGCTTCAACCCCAACA 36430
Qy	248	hrValSerAspGlyGluAsnLysLysAspLys 258
Db	36429	CAGTCTCAGATGGGAAAAACRAAGAGGATAAA 36398
RESULT 14		
AC092707		
LOCUS	AC092707	170540 bp DNA linear HTG 20-JUL-2001
DEFINITION		Homo sapiens chromosome 11 clone RP11-679G21 map 11, WORKING DRAFT
ACCESSION		SEQUENCE, 9 unordered pieces.
VERSION		AC092707
KEYWORDS		HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE		Homo sapiens.
ORGANISM		Homo sapiens
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS		1 (bases 1 to 170540)
TITLE		Homo sapiens chromosome 11, clone RP11-679G21
JOURNAL		Unpublished
REFERENCE		2 (bases 1 to 170540)
AUTHORS		Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., DeLorean, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Katat, A., Kartas, A., Kells, C., Larocque, K., Lamasares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Roman, J., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,

Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (20-JUL-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WBIR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L12706
 Center clone name: 679-G-21

----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 166430 bases at least Q40
 Consensus quality: 168426 bases at least Q30
 Consensus quality: 169078 bases at least Q20
 Insert size: 151000; agarose-fp
 Insert size: 169740; sum-of-contigs
 Quality coverage: 9.3 in Q20 bases; agarose-fp
 Quality coverage: 8.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 9 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 743: contig of 743 bp in length
 * 744 843: gap of 100 bp
 * 844 1842: contig of 999 bp in length
 * 1843 1942: gap of 100 bp
 * 1943 4117: contig of 2175 bp in length
 * 4118 4217: gap of 100 bp
 * 4218 9901: contig of 5684 bp in length
 * 9902 10001: gap of 100 bp
 * 10002 22030: contig of 12029 bp in length
 * 22031 22130: gap of 100 bp
 * 22131 77732: contig of 55602 bp in length
 * 77733 77832: gap of 100 bp
 * 77833 110109: contig of 32277 bp in length
 * 110110 110209: gap of 100 bp
 * 110210 150387: contig of 40178 bp in length
 * 150388 150487: gap of 100 bp
 * 150488 170540: contig of 20053 bp in length.

FEATURES

source
 1. 170540
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-679G21"
 /clone_lib="RPC1-11 Human Male BAC"
 misc_feature
 1. 743
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"
 844..1842
 /note="assembly_fragment"
 1943..4117
 /note="assembly_fragment"
 4218..9901
 /note="assembly_fragment"
 10002..22030

/note="assembly_fragment"
 22131..77732
 /note="assembly_fragment"
 77833..110109
 /note="assembly_fragment"
 110210..150387
 /note="assembly_fragment"
 150488..170540
 /note="assembly_fragment
 clone_end:17
 vector_side:right"
 BASE COUNT 47282 a 39323 c 37471 g 45658 t 806 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,3e-53 Length: 170540
 Score: 579.50 Matches: 172
 Percent Similarity: 27.37% Conservative: 1
 Best Local Similarity: 27.22% Mismatches: 2
 Query Match: 42.64% Indels: 459
 DB: 2 Gaps: 4
 US-09-954-846-2 (1-258) x AC092707 (1-170540)
 QY 84 MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyrPhe 103
 Db 61689 GTGTTCTCTGATGACGTGCAACCCCTATATATATGGCCCTGAGTATATCAAGTACTTC 61748
 QY 104 AsnAspLysThrIleAspGlu----- 110
 Db 61749 AATGATAAAACCATTCATGCT-GAGTCTCTTCTCCCTTCTTCTTCTGGTCCCTGTG 61807
 QY 110 ----- 110
 Db 61808 GGTGATTTTGTAGTTGTCTCTCAATTCACTAGGAGGAAACAACAGTCTTCAAAATGGAT 61867
 QY 110 ----- 110
 Db 61868 GTACGGGCACGTGTGGTTTCATTATGAGAGCCGGGGAGAGGAAACACACAGGATGAGGAAT 61927
 QY 110 ----- 110
 Db 61928 TAGATGCTAAAGTCTGAACCTTCCAGAGGCTATGTGGATCCAGCTGACTTTTCTTCCC 61987
 QY 111 ----- GluLeuGluArgAspLysArgValThrTrpIleValGluPhePhe 125
 Db 61988 TGTATTTGGCAGGAGGAGACTAGAACGGGCAAGAGGTCACCTGGATTCTGGAGTCTTT 62047
 QY 126 AlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeuSerLeuLys 145
 Db 62048 GCCAATTGGTCTAATGACTGCGCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAA- 62106
 QY 145 ----- 145
 Db 62107 GTGAGTAGTGCAAGGGAGGATGGTGGAAATGGAGATGCTGTGCTTCCCTCTCACTGT 62166
 QY 145 ----- 145
 Db 62167 TTTGGCTTTTCTTTTCTTTTGGCTTGTATTTTACACATGTTACACAAAGCATCTCC 62226
 QY 145 ----- 145
 Db 62227 CTCCTCCCTCTTAAATATCTATACTTCCACTTTCCTTGTATCCATTTTTCACGAT 62286
 QY 146 ----- TyrAsnCysThrGlyLeuAsnPhcGlyLysValAsp 157
 Db 62287 ATTAATAATATATTTCTTTTTCAGATACAACTGTACAGGCTAAATTTTGGGAAGGTGAT 62346
 QY 158 ValGlyArgTyrThrAspValSerThr----- 166
 Db 62347 GTTGACGCTATACTGATGTTAGTACGCGGTATGTAAGACCTGGCAGAGGCTCTGACC 62406
 QY 167 ----- Ar 167

```
Db 62407 AGGGAATCACTTTGAGTGATACATACAGGACATTTAGAGAACTTCTTGGGCCCTGCAG 62466
||
Qy 167 gTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeuLeuPheGlnG1 187
|||||
Db 62467 GTACAAAGTGAGCATCACCCCTCACCAAGCAACTCCCTACCTGATCCTGTGTCCAGG 62526
|||||
Qy 187 yGlyLysGluAlaMetArgArgProGlnIleAspLysLysGlyArgAlaValSerTrpTh 207
|||||
Db 62527 TGGCAAGGAGGCAATGCGCGGCCACAGATTGACAGAAAGGACGGCTGTCTCATGGAC 62586
|||||
Qy 207 rPheSerGlu----- 210
|||||
Db 62587 CTTCTGAGGTACCTGAAGGAAGGCGAGGTGCATGAAGGGTGCAGACAGTAGTGGG 62646
|||||
Qy 210 ----- 210
Db 62647 CTTTCAAGCCCTACCCGGGTTTGATTACAGCTCTGCCACTTGCCCATTAGCTTTGAGG 62706
|||||
Qy 210 ----- 210
Db 62707 TGTGGACTAGTTACTAGACCTCATGTTTAAATTCATTTAGCAAAATATATATATGTGCAG 62766
|||||
Qy 210 ----- 210
Db 62767 ACCTGCTAGTACTGGTAACAATCATATTAGATTGGTCTCAGGCCAGCAGCGTGG 62826
|||||
Qy 210 ----- 210
Db 62827 CTCACGCTGTAAATCCACGACTTTTGGGAGCGGAGCTGGCAGATCATCTGAGGTGAGG 62886
|||||
Qy 210 ----- 210
Db 62887 AGTTTGAGACCAGCCTGACCAATATGGAGAAACCCCGTCTACTAAAAATACAAAAATA 62946
|||||
Qy 210 ----- 210
Db 62947 GCCGAGTGTGGCGCATCGCTATATATGCCAGCTATTTGGGAGGCTGAGGCAGGAGAAAT 63006
|||||
Qy 210 ----- 210
Db 63007 CGCTTGAAACCGGAGGTGGAGGTTCAGTGCAGTGCAGATTGCACCATTCGCGTCCAGCC 63066
|||||
Qy 210 ----- 210
Db 63067 TGGGCAACAAGAGACTCCATCAGGAAAAAAGAAATTCGTCTCGAAGGCCACTGC 63126
|||||
Qy 210 ----- 210
Db 63127 ACCTATTGCTGTAGGTGTAGGTTCCCTGTCAGTTAATTGGAGGTAATAACCCCTGGCT 63186
|||||
Qy 210 ----- 210
Db 63187 TACAGGGTTGTTTGCAGATTACATTAATAATTTAGGTATATTTCTTGCACGGTGATT 63246
|||||
Qy 210 ----- 210
Db 63247 GGCACATAATGGGTATTTAACTATTAGTTGACTTTTCTGGGTACCTAAAAAGAGGAAGG 63306
|||||
Qy 210 ----- 210
Db 63307 TTAGGAAGATTTTGTGCTCTTGTCTTACCTTCCTTTCCAGACTTTGTGTAATAATACC 63366
|||||
Qy 210 ----- 210
Db 63367 TCTTACTTCCAGGCTCTTTACTCTCTCCCTTCCAACCCAGATCCTGACGTGTGCATCTCTT 63426
|||||
Qy 211 -----GluAsnValIleArgGluPheAsnLeuAsnGluLeuTyrGlnArgAlaLysL 228
|||||
Db 63427 TTGTGCAGGAGATGTGTATCCGAGAAATTTAACTAAATGAGCTATACCGCGGCGCAAGA 63486
|||||
Qy 228 yLeuSerLysAlaGlyAspAsnIleProGluGluGlnProValAlaSerThrProThrT 248
|||||
```

```
Db 63487 ARCTATCAAGGCTGGAGCAARTATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACCA 63546
Qy 248 hrValSerAspGlyGluAsnLysLysAspLys 258
|||||
Db 63547 CACTGTGATGGGAAACACAGAGGATAAA 63578
|||||
RESULT 15
AC021522/c
LOCUS
DEFINITION
199992 bp DNA linear HTG 21-JUL-2001
Homo sapiens clone RP11-335H5, WORKING DRAFT SEQUENCE, 14 unordered
pieces.
AC021522
AC021522.5 GI:14971311
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 199992)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Castle, A.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Coker, P.,
Choquel, Y., Collange, M., Collins, S., Collins, A., Cooke, P.,
DeArrellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardland, J., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 20, 2001 this sequence version replaced gi:14269795.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5504
Center Clone name: 335_H5
----- Summary Statistics
Sequencing vector: M13; M77815; 0% of reads
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 192697 bases at least Q40
Consensus quality: 196336 bases at least Q30
Consensus quality: 197632 bases at least Q20
Insert size: 193000; agarose-fp
Insert size: 198692; sum-of-contigs
Quality coverage: 8.3 in Q20 bases; agarose-fp
Quality coverage: 8.1 in Q20 ba.
```

* NOTE: This is a 'working draft' sequence. It currently consists of 14 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.

*	1	17413:	contig of 17413 bp in length
*	17414	17513:	gap of 100 bp
*	17514	18244:	contig of 731 bp in length
*	18245	18344:	gap of 100 bp
*	18345	19084:	contig of 740 bp in length
*	19085	19184:	gap of 100 bp
*	19185	20416:	contig of 1232 bp in length
*	20417	20516:	gap of 100 bp
*	20517	21625:	contig of 1109 bp in length
*	21626	21725:	gap of 100 bp
*	21726	102938:	contig of 81213 bp in length
*	102939	103038:	gap of 100 bp
*	103039	104905:	contig of 1867 bp in length
*	104906	105005:	gap of 100 bp
*	105006	108216:	contig of 3211 bp in length
*	108217	108316:	gap of 100 bp
*	108317	112875:	contig of 4559 bp in length
*	112876	112975:	gap of 100 bp
*	112976	122676:	contig of 9701 bp in length
*	122677	122776:	gap of 100 bp
*	122777	138821:	contig of 16045 bp in length
*	138822	138921:	gap of 100 bp
*	138922	168907:	contig of 29986 bp in length
*	168908	169007:	gap of 100 bp
*	169008	196861:	contig of 27854 bp in length
*	196862	196961:	gap of 100 bp
*	196962	199992:	contig of 3031 bp in length.

FEATURES

Location/Qualifiers
1 199992

```
1. 199992
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-335H5"
/clone_lib="RPC1-11 Human Male BAC"
```

1. .17413
/STONE_TID- RPCI-11 HUMAN MALE BAS

```

I. .I7413
/note="assembly fragment

```

```
clone_end:SP6
```

```
vector_side:left"
```

17514. 18244

```
/note="assembly_fragment"
```

18345. .19084

```
/note="assembly_fragment"
```

19185. .20416

```
/note="assembly_fragment"
```

20517. .21625

```

/note="assembly_fragment"

```

21726. .102938

```

/note="assembly_fragment"

```

103039. 104905

```

/note="assembly_fragment"
105000 100000

```

```
105006. .108216
/nota="assembly fragment"
```

```

/note="assembly_fragment"
108317 112975

```

```
108317. .112875
/note="assembly fragment"
```

```

/note="assembly_fragment"
112076 122676

```

```
II2976. .I22676
/note="assembly fragment"
```

```

/note="assembly_fragment"
122777 138821

```

```
122777. :138821
/note="assembly fragment"
```

7note= assembly_fragment
138922. . 168907

```
138922. :108907
/note="assembly fragment"
```

```

/proc/assembly_fragment
169008. 196861

```

```
105000. : 150001
/note="assembly_fragment"
```

196962. .199992

```

20002.14552
/note="assembly_fragment

```

```
clone_end:T7
```

```
vector_side:right"
```

a	46381	c	45637	g	52463	t	1305	others
---	-------	---	-------	---	-------	---	------	--------

ORIGIN

Alignment Scores:

Pred. No.:	1.59e-53	Length:	199
------------	----------	---------	-----

Percent Similarity:	27.37%	Conservative:	1
Best Local Similarity:	27.22%	Mismatches:	2
Query Match:	42.64%	Indels:	459
DB:	2	Gaps:	4
US-09-954-846-2 (1-258) x AC021522 (1-199992)			
Qy	84	MetPheLeuMetThrCysLysProProLeuTyrMetGlyProGluTyrIleLysTyrPhe	103
Db	185240	GTGTTCCCTGATGACGTGCAAAACCCCCCTATATATGCGCCTGAGTATATCAAGTACTTC	185181
Qy	104	AsnAspLysThrIleAspGlu	110
Db	185180	AATGATAAAACCATTTGATGT-GAGTGCCTTTTCCCTTTCTGTGTTCTTGGGFCCTGTG	185122
Qy	110	-----	110
Db	185121	GGTGATTTTGTAGTTGTCTCTCCATTCACTAGGAGGAACAACAGTGTCTCAAAATGGAT	185062
Qy	110	-----	110
Db	185061	GTACGGGCACTGTGGTTTCATTATGAGACCGGGGAGAGGAACACAGGATGAGGAAT	185002
Qy	110	-----	110
Db	185001	TAGATGCTAAAGTCTGAACCTTCCAGGAGGCTATGTGATCCAGCTGACTTTTCTTCCC	184942
Qy	111	-----GluLeuGluArgAspLysArgValThrTriPileValGluPhePhe	125
Db	184941	TGTAATTTGGCAGGAGGAAGTACAGACGGGACAAAGAGGTCACCTTGGATTGTGGAGTCTTT	184882
Qy	126	AlaAsnTrpSerAsnAspCysGlnSerPheAlaProIleTyrAlaAspLeuSerLeuLys	145
Db	184881	GCCAAATGGTCTAATGACTGCGCAATCATTTGCCCTATCTATGCTGAGCTCTCCCTTAA-	184823
Qy	145	-----	145
Db	184822	GTGAGTAGTGCAAAGGGAGGATGGTGGAAATGGAGATGCTGTGCTTCCCTCTCACGTGT	184763
Qy	145	-----	145
Db	184762	TTTTGGCTTTCTTTTTTGGCCTTGTATTTTCACACATGGTAACCAAGGCATCTCC	184703
Qy	145	-----	145
Db	184702	CTCTCCCTCTTAAATATCTATACTTCCACTTCTCTTGATCCATTATTTTTTTCAGCAT	184643
Qy	146	-----TyrAsnCysThrGlyLeuAsnPheGlyLysValAsp	157
Db	184642	ATTAATAATATATCTTTTCAGATACACCTGTACAGGCTTAATTTTGGGAAGGTGGAT	184583
Qy	158	ValGlyArgTyrThrAspValSerThr	166
Db	184582	GTTGGACGCTACTGATGTTAGTACGGCGTATGTAAAGACCTGGGCAGAGGGTCTGAGC	184523
Qy	167	-----Ar	167
Db	184522	AGGGAATTCATTTGAGTGATACATACAGGACATTTTAGAGAACTTTCTGGGCCCTCGAC	184463
Qy	167	gTyrLysValSerThrSerProLeuThrLysGlnLeuProThrLeuIleLeuPheGlnG1	187
Db	184462	GTACAAAGTGAGCACATCACCCCTCACCAAGCACTCCCTACCTGATCTCTGTTCCAAAG	184403
Qy	187	yGlyLysGluAlaMetArgArgProGlnIleAspLysGlyArgAlaValSerTrpTh	207
Db	184402	TGGCAAGGAGCAATGCGGGCGGCACAGATTGACAAGAAAGGCGGGCTGTCTCATGGAC	184343
Qy	207	rPheSerGlu	210
Db	184342	CTTCTGAGGTACTCTGAAAGGAAGGCGAGGTGCATGAAGGGTGCAGAACAGTAGTGGG	184283
Qy	210	-----	210

```
Db 184282 CTTTCAAGCCCTACCGGGTTTGATTACAGCTCTGCCACTTGCCCATTTAGCTTTGAGGG 184223
QY 210 ----- 210
Db 184222 TGTGGACTAGTTACTAGACCTCATGTTTAAATTCATTAGCAAAATATATATTATGTGCAG 184163
QY 210 ----- 210
Db 184162 ACCTGCTAGTACTGGGTAAACATCATTTAGTTAAGATTGGTCTCAGGCCAGGCACGGTGG 184103
QY 210 ----- 210
Db 184102 CTCACGCTGTAAATCCAGCACCTTTGGGAGCGGAGGTGGCAGATCACCTGAGGTGCAGG 184043
QY 210 ----- 210
Db 184042 AGTTTGAGACCAGCCTGACCAACATGGAGAAACCCGCTGTCTACTAAAAATACAAAAATA 183983
QY 210 ----- 210
Db 183982 GCCGAGTGTGGTGGCCCATGCCCTATAATGCCAGCTATTTGGGAGGCTGAGGCAGGAGAAT 183923
QY 210 ----- 210
Db 183922 CGCTTGAAACCGGAGGTGGAGGTTGCCAGTGAGCTGAGATTGCCACCATTCGGGTCCAGCC 183863
QY 210 ----- 210
Db 183862 TGGGCAACAAGAGACTCCATCAGGAAAAAAGAAATCTGTCTCGAAGGCCACTGC 183803
QY 210 ----- 210
Db 183802 ACCTATTGCTGTAGTGTAGTTCCTCCCTGTCAGTTAATTGGAGGTAATAAACCCTGGCT 183743
QY 210 ----- 210
Db 183742 TACAGGGTTGTTTCGAGATTACATTAATAATTTAGGTATATATCTTGGCACGGTGATT 183683
QY 210 ----- 210
Db 183682 GGCACATAATGGGTATTTAACTATTAGTTGACTTTTCTGGGTACCTAAAAAGAGGAAGG 183623
QY 210 ----- 210
Db 183622 TTAGGAAGATTTTGTCTCTTTGCTTTACTCCTTCCTTCCACACACTTTGTGTAAATACC 183563
QY 210 ----- 210
Db 183562 TCCTACTTCCAGGCTCTTTACTCTCCCTTCCAAACCCAGATCCTGAGTGTGCATCTCTT. 183503
QY 211 -----GluAsnValIleArgGluPheAsnLeuAsnGluLeuTyrGlnArgAlaLysL 228
|||||
183502 TTGTGCAGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGCCAAGA 183443
Db 228 ysLeuSerLysAlaGlyAspAsnIleProGluGluGlnProValAlaSerThrProThrT 248
|||||
183442 AACTATCAAGGCTGGAGCAATATCCCTGAGGAGCAGCCTGTGCTTCAACCCCCACCA 183383
QY 248 hrValSerAspGlyGluAsnLysLysAspLys 258
|||||
Db 183382 CAGTGTACATGGGGAACACAGAGATAAA 183351
```

Search completed: July 12, 2003, 11:49:39
Job time : 2172 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 09:18:57 ; Search time 1606 Seconds

(without alignments)
15681.184 Million cell updates/sec

Title: US-09-954-846-4

Perfect score: 1555

Sequence: 1 AGGGGAGCGGGCGGAGACC.....TAATAAAAGATTGGGATTA 1555

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

--arched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hc.*

9: gb_estl.*

10: gb_est2.*

11: gb_hc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_eston.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mam.*

24: em_gss_mus.*

25: em_gss_other.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	950.8	61.1	980	9 AL515716	AL515716 AL515716
c 2	869	55.9	1014	9 AL563867	AL563867 AL563867
3	865.4	55.7	870	14 BQ687961	BQ687961 AGENCOURT
4	865.8	55.0	1022	13 BM555220	BM555220 AGENCOURT
5	835	53.7	919	14 BQ961752	BQ961752 AGENCOURT
6	826.2	53.1	911	14 BQ940063	BQ940063 AGENCOURT

c 7	823.8	53.0	832	9	AL571100	AL571100
8	801.2	51.5	998	14	BM915245	BM915245 AGENCOURT
9	796.2	51.2	884	14	BQ430136	BQ430136 AGENCOURT
c 10	776	49.9	1017	13	B1524823	B1524823 603051831
11	765.8	49.2	838	14	BQ690091	BQ690091 AGENCOURT
12	763.2	49.1	835	14	BQ962329	BQ962329 AGENCOURT
13	751.6	48.3	913	14	BQ920164	BQ920164 AGENCOURT
14	750.8	48.3	894	14	BQ687916	BQ687916 AGENCOURT
c 15	746.2	48.0	786	9	AL569941	AL569941
16	741	47.7	1009	12	BF526515	BF526515 602070821
17	737.4	47.4	793	13	BI771888	BI771888 603055295
c 18	728.4	46.8	767	14	BM974314	BM974314 UT-CF-EC1
19	716.8	46.1	911	12	BG762562	BG762562 602734427
20	715.8	46.0	1097	14	BQ073145	BQ073145 AGENCOURT
21	713.8	45.9	827	12	BE908352	BE908352 601503159
c 22	712.2	45.8	968	12	BE748905	BE748905 601571802
23	704.4	45.3	918	14	BQ945418	BQ945418 AGENCOURT
c 24	694.2	44.6	1052	12	BF569320	BF569320 602184676
25	691.6	44.5	858	14	BQ686696	BQ686696 AGENCOURT
c 26	689.4	44.3	720	14	BM974059	BM974059 UT-CF-EC1
27	688	44.2	724	12	BE741788	BE741788 601595590
28	686.6	44.2	1008	12	BE898497	BE898497 601681402
c 29	684.4	44.0	730	14	BM988558	BM988558 UT-H-DHO-
30	684	44.0	690	14	BM762654	BM762654 K-EST0043
c 31	683.4	43.9	740	14	BM975449	BM975449 UT-CF-EN1
32	682	43.9	834	12	BG179545	BG179545 602328142
c 33	681.2	43.8	702	9	AL570762	AL570762
34	681.2	43.8	948	12	BE796785	BE796785 601587565
35	679.8	43.7	880	10	BE616524	BE616524 601281171
36	679	43.7	860	13	B1161293	B1161293 602865578
37	678.6	43.6	928	12	BG251786	BG251786 602363694
38	677.8	43.6	1142	13	BM559140	BM559140 AGENCOURT
c 39	676.6	43.5	710	14	BQ447137	BQ447137 UI-H-EU1-
40	673.6	43.3	766	12	BG762793	BG762793 602734719
c 41	671.4	43.2	698	14	BQ009303	BQ009303 UI-H-ED1-
c 42	670.4	43.1	731	14	BM975356	BM975356 UI-CF-EN1
c 43	669.6	43.1	779	14	BM981327	BM981327 UI-CF-EN1
44	669.4	43.0	838	12	BF342782	BF342782 602015055
c 45	662.2	42.6	716	14	BM978725	BM978725 UI-CF-DU1

ALIGNMENTS

RESULT 1
AL515716/c
LOCUS AL515716 LTI_NFL011_NBC1 980 bp mRNA linear EST 13-FEB-2001
DEFINITION Prime, mRNA sequence.
ACCESSION AL515716
VERSION AL515716.1 GI:12779209
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 980)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Location/Qualifiers
1..980
Source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODA001YG10"
/clone_lib="LTI_NFL011_NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"

/note="Organ: brain; Vector: pcMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcMVSPORT 6 vector. Library is not normalized, but is the control for the normalized libraries. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 264 a 229 c 229 g 253 t 5 others
ORIGIN

Query Match 61.1%; Score 950.8; DB 9; Length 980;
Best Local Similarity 98.8%; Pred. No. 2.4e-286;
Matches 953; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 554 ACACGCTACAAAGTGAGCATCACCCCTCACCAGCAACTCCCTACCTGATCCTGTC 613
DB 974 ACCGGGTACAAAGTGAGCACTCACCCCTCACCAGCAACTCCCTACCTGATCCTGTC 915
614 CRAAGTGGCAAGGAGGCAATGCGCGGCCACAGATTGACAAAGAGGAGCGGCTGTCTCA 673
914 CRAAGTGGCAAGGAGGCAATGCGCGGCCACAGATTGACAAAGAGGAGCGGCTGTCTCA 855
QY 674 TGGACCTTCTCTGAGGAGATGTATCCGAGATTTAACTTAATGAGCTATACGAGGG 733
DB 854 TGGACCTTCTCTGAGGAGATGTATCCGAGATTTAACTTAATGAGCTATACGAGGG 795
QY 734 GCCAAGAAATATCAAGGCTGAGACAATATCCTCAGGAGCAGCTGTGGCTTCAACC 793
DB 794 GCCAAGAAATATCAAGGCTGAGACAATATCCTCAGGAGCAGCTGTGGCTTCAACC 735
QY 794 CCACACACAGTGCAGATGGGAAAACAAGAGGATATAAAGATCTCTCACTTTGGCAGT 853
DB 734 CCACACACAGTGCAGATGGGAAAACAAGAGGATATAAAGATCTCTCACTTTGGCAGT 675
QY 854 GCTTCCTCTCTGATTAATCCAGCTCTTTCATTAACCAAGCCTGAGGCTGAGGCTT 913
DB 674 GCTTCCTCTCTGATTAATCCAGCTCTTTCATTAACCAAGCCTGAGGCTGAGGCTT 615
QY 914 TTATTATGTTTCCCTTTGGCTGTGACTGGTGGGCGACATCGACTCTGATTTAA 973
DB 614 TTATTATGTTTCCCTTTGGCTGTGACTGGTGGGCGACATCGACTCTGATTTAA 555
QY 974 AGAGGCATCTAGGAAATGTCAGGCACCCCTACAGGAAGGCTGCGATGCTGTGGCCAACT 1033
DB 554 AGAGGCATCTAGGAAATGTCAGGCACCCCTACAGGAAGGCTGCGATGCTGTGGCCAACT 495
QY 1034 GTTTCCTGAGCAAGAAAGAGATCTCATAGGAGGAGGGGAAATGTTTCCCTCCAAAG 1093
DB 494 GTTTCCTGAGCAAGAAAGAGATCTCATAGGAGGAGGGGAAATGTTTCCCTCCAAAG 435
QY 1094 CTTGGGTGAGTGTGTTAACTGCTTATCAGCTATTTCAGACATCTCCATGTTTCCCAATGA 1153
DB 434 CTTGGGTGAGTGTGTTAACTGCTTATCAGCTATTTCAGACATCTCCATGTTTCCCAATGA 375
QY 1154 AACTCTGTGTTTCATCATCTCTTCTAGTTGACCTGCACAGCTGTGTTAGACCTAGATT 1213
DB 374 AACTCTGTGTTTCATCATCTCTTCTAGTTGACCTGCACAGCTGTGTTAGACCTAGATT 315
QY 1214 TAACCTTAAGGTAAAGTGTGTTGATAGAACGCTTAAGAAATTTTCCCAAGGACTCTTG 1273
DB 314 TAACCTTAAGGTAAAGTGTGTTGATAGAACGCTTAAGAAATTTTCCCAAGGACTCTTG 255
QY 1274 CTTCCCTTAAGCCCTCTGCTGCTTTATGCTTCAATTAAGATATAAGCTTAACCTTTG 1333
DB 254 CTTCCCTTAAGCCCTCTGCTGCTTTATGCTTCAATTAAGATATAAGCTTAACCTTTG 195
QY 1334 TCCTAGTCTTAAGGAGAAACCTTTTAACCAACAAAGTTTTATCATTTGAAGACAAATTTGA 1393
DB 194 TCCTAGTCTTAAGGAGAAACCTTTTAACCAACAAAGTTTTATCATTTGAAGACAAATTTGA 135

QY 1394 ACACCCCTATTTTGGGGATTGAGAGGGTGAATAGAGGCTTGAGACTTTCCCTTTG 1453
DB 134 ACACCCCTATTTTGGGGATTGAGAGGGTGAATAGAGGCTTGAGACTTTCCCTTTG 75
QY 1454 TGTGGTAGGACTTGGAGGAGAAATCCCTCTGACATTTTCACTAACCTCTTGACATACTCCC 1513
DB 74 TGTGGTAGGACTTGGAGGAGAAATCCCTCTGACATTTTCACTAACCTCTTGACATACTCCC 15
QY 1514 ACAC 1518
DB 14 GCAVC 10

RESULT 2
AL563867/c
LOCUS
DEFINITION AL563867 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD008YG11 3
prime, mRNA sequence.
ACCESSION AL563867
VERSION AL563867.1 GI:12913683
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Location/Qualifiers
1..1014
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DD008YG11"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pcMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 277 a 230 c 229 g 248 t 30 others
ORIGIN

Query Match 55.9%; Score 869; DB 9; Length 1014;
Best Local Similarity 91.4%; Pred. No. 1.1e-260;
Matches 909; Conservative 24; Mismatches 60; Indels 2; Gaps 2;

QY 486 CCCTTAATCAACTGTACAGGCTAAATTTTGGGAAGGTGATGTTGGAGCGCTACTG 545
DB 1007 CCTTAAAAACACTGACAGGCTAAATTTTGGGAAGGTGATGTTGGAGCGCTAAACTG 948
QY 546 ATGTTAGTAGCGGTGACAAAGTGAGCATCACCCCTCACCAAGCAACTCCCTACCTGA 605
DB 947 ATGTTGGAACGCGTACAAAGGAGACWTACCCCTCMCCARGCAACTCCCTACCTGA 888
QY 606 TCCGTGTTCCAAAGTGGCAAGGAGCAATGCGGCGCCACAGATTGACAAAGAAAGACGGG 665
DB 887 TCCTGTCCCAARGTGGCAAGGAGCAATGCGGCGCCACCAATTAACAAAGAAAGACGGG 828
QY 666 CTGTCTCATGACCTCTCTCTGAGGAGATGTGATCCGAGAAATTTAACTTAATGAGCTAT 725

```
Db 827 YTTTTCATGACCTCTCTGAGGARAATKTGATCCGAGAATTTAACTTAAATGAGCTAT 768
Qy 726 ACCAGCGGGCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGAGAGCCTGTGG 785
Db 767 ACCGCGGGCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGCGAGCTGTGG 708
Qy 786 CTTCAACCCCCACACAGCTGTGAGATGGGGAACAAGAGGATAAATAAGATCTCTACT 845
Db 707 CTTCAACCCCCACACAKTKTCAGATGGGGAACAAGAGGATAAATAAGACCTTACT 648
Qy 846 TTGGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 905
Db 647 TTGGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 588
Qy 906 GCAGCCTTTTATTTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 965
Db 587 GCAGCCTTTTATTTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 528
Qy 966 GATTTTAAAGAGGATCTAGGGAATTTGTAGGACCCCTACAGGAAGGCTGCCATGCT 1025
Db 527 GATTTTAAAGAGGATCTAGGGAATTTGTAGGACCCCTACAGGAAGGCTGCCATGCT 468
Qy 1026 GGCAACTGTTTCTACTGGAGCAAGAAAGATCTATAGGACGGGAGGGAATGTTTC 1085
Db 467 GGCCCACTGTTTCTACTGGAGCAAGAAAGGCTCTCATAGGACGGGAGGGAATGTTTC 408
Qy 1086 CTTCAAGCTTGGTCTAGTGTAACTGCTTATCAGCTATTCAGACATCTCCATGGTTT 1145
Db 407 CTTCAAGCTTGGTCTAGTGTAACTGCTTATCAGCTATTCAGACATCTCCATGGTTT 349
Qy 1146 CTCCATGAACCTCTGTGTTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1205
Db 348 CTCCATGAACCTCTGTGTTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 289
Qy 1206 CCTAGATTTAACCTTAAGGTAAAGTCTGTGGGTATAGAACGCTAAGAAATTTTCCCAAG 1265
Db 288 CCTAGATTTAACCTTAAGGTAAAGTCTGTGGGTATAGAACGCTAAGAAATTTTCCCAAG 229
Qy 1266 GACTCTTCTCTCTTAAAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1325
Db 228 AACTCTGCTCTCTTAAAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 169
Qy 1326 TAACCTTCTCTCTTAAAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1385
Db 168 TAACCTTCTCTCTTAAAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 109
Qy 1386 AATATTGAACAACCCCTATTTTCTGGGATTTGAGAGGGGTGAATAGAGCCTTGAGACT 1445
Db 108 AATATTGAACAACCCCTATTTTCTGGGATTTGAGAGGGGTGAATAGAGCCTTGAGACT 49
Qy 1445 TTCTTTTCTGTGTAGGACTTGGAGGAGAAATCCC 1480
Db 48 TTCTTTTCTGTGTAGGACTTGGAGGAGAAATCCC 15
```

```
RESULT 3
B0687961
LOCUS
DEFINITION B0687961 870 bp mRNA linear EST 15-JUL-2002
5', mRNA sequence.
ACCESSION B0687961
VERSION B0687961.1 GI:21813277
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 870)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
```

```
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM2388 row: o column: 05
High quality sequence stop: 647.
Location/Qualifiers
1..870
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_110"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pMTB7; Site:1: XhoI;
Site:2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 227 a 203 c 215 g 224 t 1 others
ORIGIN
Query Match 55.7%; Score 866.4; DB 14; Length 870;
Best Local Similarity 99.8%; Pred. No. 6.8e-260;
Matches 867; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 369 ATGATAAAACCATTTGATGAGGAACATAGAACGGGACAGAGGCTCACTTGGATTGGAGCT 428
Db 1 ATGATAAAACCATTTGATGAGGAACATAGAACGGGACAGAGGCTCACTTGGATTGGAGCT 60
Qy 429 TCTTTGCCAATTTGCTTAATGACTGCCAATCATTTTGGCCCTATCTATGCTGACCTCTCC 488
Db 61 TCTTTGCCAATTTGCTTAATGACTGCCAATCATTTTGGCCCTATCTATGCTGACCTCTCC 120
Qy 489 TTAATAACAACCTGTACAGGGCTAAATTTTGGGAAGTGGATGTTGGACCCCTACTACTGATG 548
Db 121 TTAATAACAACCTGTACAGGGCTAAATTTTGGGAAGTGGATGTTGGACCCCTACTACTGATG 180
Qy 549 TTAGTACCGGTTACAAAGTACAGCATCACCCCTCAACGAAGCACTCCCTACCCCTGATCC 608
Db 181 TTAGTACCGGTTACAAAGTACAGCATCACCCCTCAACGAAGCACTCCCTACCCCTGATCC 240
Qy 609 TGTTCGAAGTGGCAAGGAGGCAATGCGGGGCCACAGATTGACAAGAAAGGACGGGTG 668
Db 241 TGTTCGAAGTGGCAAGGAGGCAATGCGGGGCCACAGATTGACAAGAAAGGACGGGTG 300
Qy 669 TCTCATGACCTCTCTCTGAGGAGATGTATCCCGAGAATTTAACTTAATGAGCTATACC 728
Db 301 TCTCATGACCTCTCTCTGAGGAGATGTATCCCGAGAATTTAACTTAATGAGCTATACC 360
Qy 729 AGCGGGCCCAAGAACTATCAAAAGGCTGGAGACAATATCCCTGAGGACGCTGTGGT 788
Db 361 AGCGGGCCCAAGAACTATCAAAAGGCTGGAGACAATATCCCTGAGGACGCTGTGGT 420
Qy 789 CAACCCCCACACAGTGTCTAGATGGGGAACAAGAGGATAAATAAGATCTCTACTTTG 848
Db 421 CAACCCCCACACAGTGTCTAGATGGGGAACAAGAGGATAAATAAGATCTCTACTTTG 480
Qy 849 GCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 908
Db 481 GCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Qy 909 GCCTTTTATTTATGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 968
Db 541 GCCTTTTATTTATGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
```

```
QY 969 TTTAAGAGGCATCTAGGAATTTGTCAGGACCCCTACAGAAAGGCTGCGCATGCTGGC 1028
Db 601 TTTAAGAGGCATCTAGGAATTTGTCAGGACCCCTACAGAAAGGCTGCGCATGCTGGC 660
QY 1029 CAACGTGTTTCACTGGAGCAAGAAAGAGATCTCATAGGACGGGGGGAATGTTTCCCT 1088
Db 661 CAACGTGTTTCACTGGAGCAAGAAAGAGATCTCATAGGACGGGGGGAATGTTTCCCT 720
QY 1089 CCAAGCTTGGGTCAGTGTGTTTAACTGCTTATCAGCTATTACAGACATCTCCATGTTTCTC 1148
Db 721 CCAAGCTTGGGTCAGTGTGTTTAACTGCTTATCAGCTATTACAGACATCTCCATGTTTCTC 780
QY 1149 CATGAACCTCTCTGTTTTCATCTCTTCTAGTTGACCTGCACAGCTTGGTTAGACCT 1208
Db 781 CATGAACCTCTCTGTTTTCATCTCTTCTAGTTGACCTGCACAGCTTGGTTAGACCT 840
QY 1209 AGATTAAACCTTAAGTAAAGTAAAGTCTGGG 1237
Db 841 AGATTAAACCTTAAGTAAAGTAAAGTCTGGG 869

RESULT 4
- "55220
JS
- INITIATION
AGENCY 6541254 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5550270
5' mRNA sequence.
ACCESSION
BM555220
VERSION
BM555220.1 GI:18795504
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1022)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12262 row: c column: 07
High quality sequence stop: 726.
FEATURES
Location/Qualifiers
1..1022
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5550270"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC library."
BASE COUNT
246 a 261 c 255 g 251 t 9 others
ORIGIN
Query Match 55.0%; Score 855.8; DB 13; Length 1022;
Best Local Similarity 99.1%; Pred. No. 1.5e-256;
Matches 871; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 62 GCGGCTTGGCACCTTAATGCTCTGCTGATTCGTCGCGGACTTTCACCATGGCTC 121
Db 1 GCGGCTTGGCACCTTAATGCTCTGCTGATTCGTCGCGGACTTTCACCATGGCTC 60
QY 122 GCCCAACCTTACTACCTTCTGTCGGGCCCTGCTCTGCTGCTTCTACTGTCGAGAA 181
```

```
Db 61 GCCCAACCTTACTACCTTCTGTCGGGCCCTGCTCTGCTGCTTCTACTCTGTCGAGAA 120
QY 182 CTGCCGCCCTCTGCCACGGTCTGCCACCCCAACCGGAAAGACGGTAAACCGTGTGACTTT 241
Db 121 CTGCCGCCCTCTGCCACGGTCTGCCACCCCAACCGGAAAGACGGTAAACCGTGTGACTTT 180
QY 242 GACTGAGAGAAGTGGAGATCTTCTCAGTGCCTTGTGATGATGAGAACCGC 301
Db 181 GACTGAGAGAAGTGGAGATCTTCTCAGTGCCTTGTGATGATGAGAACCGC 240
QY 302 AGATCATGTTCTCTGATGACGTGCAAAACCCCTATATATATGAGTGCCTTGTGATGATCAAG 361
Db 241 AGATCATGTTCTCTGATGACGTGCAAAACCCCTATATATATGAGTGCCTTGTGATGATCAAG 300
QY 362 TACTTCAATGATAAAACCAATTTGATGAGGAATAGAACGGGACAGAGGTCCTTGGATT 421
Db 301 TACTTCAATGATAAAACCAATTTGATGAGGAATAGAACGGGACAGAGGTCCTTGGATT 360
QY 422 GTGGAGTCTTTGCCAATTTGGTCTAATGACTGCAATCATTTGCCCTATCTATGCTGAC 481
Db 361 GTGGAGTCTTTGCCAATTTGGTCTAATGACTGCAATCATTTGCCCTATCTATGCTGAC 420
QY 482 CTCTCCCTTAAATACAACTGTACAGGGCTAAATTTTGGGAAAGTGGATGTTGGACGCTAT 541
Db 421 CTCTCCCTTAAATACAACTGTACAGGGCTAAATTTTGGGAAAGTGGATGTTGGACGCTAT 480
QY 542 ACTGATGTTAGTAGCGGGTACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 601
Db 481 ACTGATGTTAGTAGCGGGTACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 540
QY 602 CTGATCTGTTTCAAGGTGCAAGGAGGCAATGGGGGGGACAGATGTCACAGAAAGGA 661
Db 541 CTGATCTGTTTCAAGGTGCAAGGAGGCAATGGGGGGGACAGATGTCACAGAAAGGA 600
QY 662 CGGGCTGCTCATGAGACCTTCTCTGAGGAGAAATGTGATCCGAGAAATTAACCTTAAATGAG 721
Db 601 CGGGCTGCTCATGAGACCTTCTCTGAGGAGAAATGTGATCCGAGAAATTAACCTTAAATGAG 660
QY 722 CTATACAGCGGGCCCAAGAACTATCAAGGCTGGAGACAAATATCCCTGAGGACGCT 781
Db 661 CTATACAGCGGGCCCAAGAACTATCAAGGCTGGAGACAAATATCCCTGAGGACGCT 720
QY 782 GTGGCTTCAACCCCGGACAGTGTGATGAGGAGAAAGAGAGATTAATAGATCCT 841
Db 721 GTGGCTTCAACCCCGGACAGTGTGATGAGGAGAAAGAGAGATTAATAGATCCT 780
QY 842 CACTTTGGCAGTCTTCTCTCTCTCAATTCAGGCTCTTCCATACACACAGCCTGA 901
Db 781 CACTTTGGCAGTCTTCTCTCTCTCAATTCAGGCTCTTCCATACACACAGCCTGA 840
QY 902 GGCTGCAG-CCTTTTATTTTCCCTTTCCCTTTGGCTGTG 939
Db 841 AGCTGCAGNCTTTTATTTATTTTCCCTTTGGCCTG 879

RESULT 5
BQ961752
LOCUS
BO961752
DEFINITION
AGENCOURT_8820550 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:6379836
5' mRNA sequence.
ACCESSION
BQ961752
VERSION
BO961752.1 GI:22377230
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 919)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2567 row: h column: 13
High quality sequence stop: 727.
Location/Qualifiers
1. .919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6379836"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library." 8 others
BASE COUNT 223 a 238 c 222 g 228 t
ORIGIN

Query Match 53.7%; Score 835; DB 14; Length 919;
Best Local Similarity 97.8%; Pred. No. 4.9e-250;
Matches 874; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 49 GCGCGAAGATGCGGCTTGGACCTCTAATGCTCTCGTATTCGGTGGCGCGACT 108
DB 13 GGGCCGAAGATGCGGCTTGGACCTCTAATGCTCTCGTATTCGGTGGCGCGACT 72
QY 109 TTACAGATGGCTGCCCAACCTTACTACTCTCTCGGCGCTCTCTCTGCTGCTTCTCT 168
DB 73 TTACAGATGGCTGCCCAACCTTACTACTCTCTCGGCGCTCTCTCTGCTGCTTCTCT 132
QY 169 ACTGTTGAGAACTGCGCGCTCTGCCAGGCTCTGCCACCCAGCGAAGAGCTAA 228
DB 133 ACTGTTGAGAACTGCGCGCTCTGCCAGGCTCTGCCACCCAGCGAAGAGCTAA 192
QY 229 CCGTGTGACTTTGACTGGAGAGAGTGGAGATCCTGATGTTCTCAGTGCATTTGAT 288
DB 193 CCGTGTGACTTTGACTGGAGAGAGTGGAGATCCTGATGTTCTCAGTGCATTTGAT 252
QY 289 GATGAAGAACCGAGATCCATGTCCTGATGACGTGCAACCCCTATATATGGGCC 348
DB 253 GATGAAGAACCGAGATCCATGTCCTGATGACGTGCAACCCCTATATATGGGCC 312
QY 349 TGAGTATATCAAGTACTTCAATGATAAACCATTGATGAGGAAGTACAGCGGACAAG 408
DB 313 TGAGTATATCAAGTACTTCAATGATAAACCATTGATGAGGAAGTACAGCGGACAAG 372
QY 409 GTTCACTTGGATTGTGGAGTCTTTGGCAATGGTCTTAATGACTGCCAATCATTTGCC 468
DB 373 GTTCACTTGGATTGTGGAGTCTTTGGCAATGGTCTTAATGACTGCCAATCATTTGCC 432
QY 469 TATCTATGCTGACCTCTCCCTTAATCAACTGACAGGCTTAATTTGGGAGGTGGA 528
DB 433 TATCTATGCTGACCTCTCCCTTAATCAACTGACAGGCTTAATTTGGGAGGTGGA 492
QY 529 TGTGGACGCTATACCTGATGTTAGTACGCGGTAAAGTGGACATACCCCTCACC 588
DB 493 TGTGGACGCTATACCTGATGTTAGTACGCGGTAAAGTGGACATACCCCTCACC 552
QY 589 GCAACTCCCTACCTGATCTGTTCCAAAGTGGCAAGGAGGCAATGCGGCGGCACAG 648
DB 553 GCAACTCCCTACCTGATCTGTTCCAAAGTGGCAAGGAGGCAATGCGGCGGCACAG 612

649 TGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTGAGGAGAATGTGATCCGAGAATT 708
DB 613 TGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTGAGGAGAATGTGATCCGAGAATT 672
QY 709 TAACTTAATAGTATATACCAAGCGGCGCAAGAACTATCAAAAGGCTGGAGACAATATCCC 768
DB 673 TAACTTAATAGTATATACCAAGCGGCGCAAGAACTATCAAAAGGCTGGAGACAATATCCC 732
QY 769 TGAGGAGACGCTGTGGCTTCAACCCACCACAGAGTGTGAGATGGGGAACAAAGGA 828
DB 733 TGAGGAGACGCTGTGGCTTCAACCCACCACAGAGTGTGAGATGGGGAACAAAGGA 792
QY 829 TAAATAAGATCTCACTTTGGCAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 887
DB 793 TAAATAAGATCTCACTTTGGCAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 852
QY 888 AACCAAC-AAGCCTGAGGCTGCAGCCCTTTT-ATTATGTTTTCCTTTTGGCTGTG 939
DB 853 TACCACAAACCTGAGGCTGCAGCCCTTTTAAATATGTTTTCCTTTTGGCTGTG 906

RESULT 6
BQ940063
LOCUS
DEFINITION
AGENCOURT_8865577 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6341656
5', mRNA sequence.
ACCESSION BQ940063
VERSION BQ940063.1 GI:22355541
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 911)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2540 row: a column: 17
High quality sequence stop: 727.
Location/Qualifiers
1. .911
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6341656"
/clone_lib="NIH_MGC_47"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library." 1 others
BASE COUNT 236 a 206 c 223 g 245 t
ORIGIN

Query Match 53.1%; Score 826.2; DB 14; Length 911;
Best Local Similarity 97.0%; Pred. No. 2.8e-247;
Matches 875; Conservative 0; Mismatches 19; Indels 8; Gaps 3;

```
QY 588 AGCAACTCCCTACCTGATCCTGTTCACAGGTGGCAGGAGGCAATGCGGCGCCACAGA 647
Db 1 AGCAACTCCCTACCTGATCCTGTTCACAGGTGGCAGGAGGCAATGCGGCGCCACAGA 60
QY 648 TTGACAAAGAAAGCAGCGGTGCTCTGATGACCTTCTGAGGAGCAATGTGATCGGAGAAT 707
Db 61 TTGACAAAGAAAGCAGCGGTGCTCTGATGACCTTCTGAGGAGCAATGTGATCGGAGAAT 120
QY 708 TTAACCTTAATGAGCTATACCAAGCGGCCCAAGAACTATCAAAAGGCTGGAGACAATATCC 767
Db 121 TTAACCTTAATGAGCTATACCAAGCGGCCCAAGAACTATCAAAAGGCTGGAGACAATATCC 180
QY 768 CTGAGAGCAGCCTGTGGCTTCAACCCCAACACAGCTGTCAGATGGGAAACAGAGG 827
Db 181 CTGAGAGCAGCCTGTGGCTTCAACCCCAACACAGCTGTCAGATGGGAAACAGAGG 240
QY 828 ATAAATAAGATCCTCACTTTGGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 887
Db 241 ATAAATAAGATCCTCACTTTGGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 888 AACCACAAGCCTGAGGCTGAGCCTTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 947
Db 301 AACCACAAGCCTGAGGCTGAGCCTTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 360
QY 948 GGCACATGCGACCTTCTGATTTTAAAGAGGCATCTAGGAAATGTGACGACCCCTACAG 1007
Db 361 GGCACATGCGACCTTCTGATTTTAAAGAGGCATCTAGGAAATGTGACGACCCCTACAG 420
QY 1008 GAAGGCTGCGATGCTGTGGCCCAACTGTTTCACTGGAGCAAGAAAGATCTCATAGGAC 1067
Db 421 GAAGGCTGCGATGCTGTGGCCCAACTGTTTCACTGGAGCAAGAAAGATCTCATAGGAC 480
QY 1068 GGAGGGGGAATGTTTCCCTCCCAAGCTGGGTCAGTGTGTTTAACTGCTTATCATAGCTATT 1127
Db 481 GGAGGGGGAATGTTTCCCTCCCAAGCTGGGTCAGTGTGTTTAACTGCTTATCATAGCTATT 540
QY 1128 CAGACATCTCCATGGTTTCTCCATGAACCTGTGTTTTCATCATCTCTTCTTCTTCTTCTT 1187
Db 541 CAGACATCTCCATGGTTTCTCCATGAACCTGTGTTTTCATCATCTCTTCTTCTTCTTCTT 600
QY 1188 CTGACAGCTTGGTTAGACCTAGATTTTAAACCCCTAAGCTAGTGTGGGTATAGAACGC 1247
Db 601 CTGACAGCTTGGTTAGACCTAGATTTTAAACCCCTAAGCTAGTGTGGGTATAGAACGC 660
QY 1248 TAAGAAATTTTCCCAAGGACTCTTGCTTCCCTTAAGCCCTTCTGGCTTCTGTTTATGCTCT 1307
Db 661 TAAGAAATTTTCCCAAGGACTCTTGCTTCCCTTAAGCCCTTCTGGCTTCTGTTTATGCTCT 720
QY 1308 TCATTAAGATATAAGCCTTAACCTTGTGCTAGTCTTAAAGGAGAAACCTTTAAACCAAAA 1367
Db 721 TCATTAAGATATAAGCCTTAACCTTGTGCTAGTCTTAAAGGAGAAACCTTTAAACCAAAA 780
QY 1368 GTTTTATCATTTGAAGACAATATTGAACACCCCTCTTATTTGTTGGGATTTGAGAAGG 1424
Db 781 GTTTTATCATTTGAAGACAATATTGAACACCCCTCTTATTTGTTGGGATTTGAGAAGG 840
QY 1425 GGTGAATAGA--GGCTTGAGACTTTTCC--TTGTTGTGTGAGACTTTGAGGAGGAATATCC 1479
Db 841 GGTGAATAGAAGGCTTGAGACTTTTCCCTTTGTTGTGTGAGACTTTGAGGAGGAGGAAT 900
QY 1480 CC 1481
Db 901 CC 902
```

RESULT 7

```
AL571100/c 832 bp mRNA linear EST 16-FEB-2001
LOCUS AL571100 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1027YB05 3
DEFINITION prime, mRNA sequence.
ACCESSION AL571100
VERSION AL571100.1 GI:12928060
KEYWORDS EST.
```

SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 832)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source

1..832
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1027YB05"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pcMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pcMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 236 a 188 c 191 g 214 t 3 others

ORIGIN
Query Match 53.0%; Score 823.8; DB 9; Length 832;
Best Local Similarity 99.2%; Pred. No. 1.5e-246;
Matches 825; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 702 GAGATTTAACTTAATGAGCTATACCAAGCGGCCCAAGAACTATCAAGGCTGGAGACA 761
Db 832 GAGATTTAACTTAATGAGCTATACCAAGCGGCCCAAGAACTATCAAGGCTGGAGACA 773
QY 762 ATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCAACACAGTGTGAGATGGGAAACA 821
Db 772 ATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCAACACAGTGTGAGATGGGAAACA 713
QY 822 AGAAGATAAATAAGATCTCAGCTTTGGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 881
Db 712 AGAAGATAAATAAGATCTCAGCTTTGGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 653
QY 882 TTCCATACCAACAGCCTGAGGCTGAGCCTTTTATTTATTTATTTATTTATTTATTTATTT 941
Db 652 TTCCATACCAACAGCCTGAGGCTGAGCCTTTTATTTATTTATTTATTTATTTATTTATTT 593
QY 942 TGGTGGGCGGAGCATGAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTTCTCAGGCACC 1001
Db 592 TGGTGGGCGGAGCATGAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTTCTCAGGCACC 533
QY 1002 CTACAGGAAGCCTGCGCATGCTGTGGCCAACTGTTTCACTGGACAGAAAGAGATCTCA 1061
Db 532 CTACAGGAAGCCTGCGCATGCTGTGGCCAACTGTTTCACTGGACAGAAAGAGATCTCA 473
QY 1062 TAGGACGGAGGGGAAATGGTTTCCCTCCCAAGCTTGGTCACTGTGTAAGTCTTCTTATCA 1121
Db 472 TAGGACGGAGGGGAAATGGTTTCCCTCCCAAGCTTGGTCACTGTGTAAGTCTTCTTATCA 413
QY 1122 GCTATTCAGACATCTCCATGGTTTCTCCATGAACCTCTGGTGTTCATCTTCTCTTA 1181
Db 412 GCTATTCAGACATCTCCATGGTTTCTCCATGAACCTCTGGTGTTCATCTTCTCTTA 353
QY 1182 GTTGACCTGCACAGCTTGGTTAGACCTTAGATTTAAACCCCTAAGGTAAGATGCTGGGGTATA 1241
Db 352 GTTGACCTGCACAGCTTGGTTAGACCTTAGATTTAAACCCCTAAGGTAAGATGCTGGGGTATA 293
QY 1242 GAAGCTTAAGAAATTTTCCCAAGGACTCTTGTCTTCTTAAAGCCCTTCTGGCTTCTGTTA 1301

Db 956 GGGAGTAATACCCCTTCCCAAGCGAGCTTCCCTTACCGGACCCGCTGTTCCCAACGGTGGC 897
 Qy 622 CAAGGAGGCAATCGCGCGGCCA-CAGATTGACAAGAAAGGAC---GGGCTGTCTCATGGA 677
 Db 896 AAGGAGCGCAATCGCGCGGCCAGCAGATGACAGAAAGAGCCGGCTGGTCTCAGGGA 837
 Qy 678 CTTCTCTGAGGAGATGTATCGGAGAAATTAACATTAATAGCTATACCAAGCGGCGCA 737
 Db 836 CTTCTCTGAGAGAAATGTATCGGAGAAATTAACATTAATAGCTATACCAAGCGGCGCA 777
 Qy 738 AGAACTATCAAGGCTGGAGACAATATCCCTGA-CGAGCAGCCTGTGCTTCAACCCCC 796
 Db 776 GAACCTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGCTTCAACCCCC 717
 Qy 797 ACCACAGTGTAGATCGGGAACAGAGGATAAATAAGATCTCTCACCTTTGGCAGTGT 856
 Db 716 ACCACAGTGTAGATCGGGAACAGAGGATAAATAAGATCTCTCACCTTTGGCAGTGT 657
 Qy 857 TCCTCTCTGCAATTCAGGCTCTTTCCAT-AACACAGCCTGAGGCTGCAGCCTTTT 915
 Db 656 TCCTCTCTGCAATTCAGGCTCTTTCCATGAACACAGCCTGAGGCTGCAGCCTTTT 597
 Qy 916 ATTTATGTTTCCCTTGGCTGACTGGTGGGCGAGCATGCAGCTTCTGATTTAAAG 975
 Db 596 ATTTATGTTTCCCTTGGCTGACTGGTGGGCGAGCATGCAGCTTCTGATTTAAAG 537
 Qy 976 AGGCATCTAGGAAATGTGAGCACCCTACAGAAAGGCTGCATGCTGTGGCCAACTGT 1035
 Db 536 AGGCATCTAGGAAATGTGAGCACCCTACAGAAAGGCTGCATGCTGTGGCCAACTGT 477
 Qy 1036 TTCCTGAGCAAGAAAGAGATCTCATAGGACGAGGGGAAATGTTTCCCTTCCAAAGT 1095
 Db 476 TTCCTGAGCAAGAAAGAGATCTCATAGGACGAGGGGAAATGTTTCCCTTCCAAAGT 417
 Qy 1096 TGGTCACTGTGTTAACTGCTTATCAGCTATTCAGACATCTCCATGTTTCTCCATGAA 1155
 Db 416 TGGTCACTGTGTTAACTGCTTATCAGCTATTCAGACATCTCCATGTTTCTCCATGAA 357
 Qy 1156 CTCTGTGTTTTCATCATCTCTTCTAGTTGACCTGCACAGCTTGTGTAGACCTAGATTTA 1215
 Db 356 CTCTGTGTTTTCATCATCTCTTCTAGTTGACCTGCACAGCTTGTGTAGACCTAGATTTA 297
 Qy 1216 ACCCTAAGTAAAGTGTCTGGGTATAGAACCGCTAAGAAATTTTCCCCCAAGGACTCTTGT 1275
 Db 296 ACCCTAAGTAAAGTGTCTGGGTATAGAACCGCTAAGAAATTTTCCCCCAAGGACTCTTGT 237
 Qy 1276 TCCTTAAGCCCTTCTGGCTTCTGTTTATGCTTCA-TTAAAGTATAAGCCCTTAACCTTGT 1334
 Db 236 TCCTTAAGCCCTTCTGGCTTCTGTTTATGCTTCACTTAAAGTATAAGCCCTTAACCTTGT 177
 Qy 1335 CGCTAGTCTAAGGAGAAACCTTTAAACCAAAAGTTTTTATCATTAAGACAAATATGAA 1394
 Db 176 CGCTAGTCTAAGGAGAAACCTTTAAACCAAAAGTTTTTATCATTAAGACAAATATGAA 117
 Qy 1395 CAACCCCTATTTTGTGGGATGAGAA--GGGTGAATAGAGGCTTGGAGCTTTCCCTT 1452
 Db 116 CAACCCCTATTTTGTGGGATGAGAAACCGGGTGAATAGAGGCTTGGAGCTTTCCCTT 57
 Qy 1453 GTGTGGTAGGACTTGGAGGAGAAATCCCTGGACTTTTCACTAACCCCTTGACATCTC 1510
 Db 56 GTGTGGTAGGACTTGGAGGAGAA--TCCCTCTGGACTTCAATAAACCCCTTGACATCTC 1

RESULT 11

B0690091
 LOCUS B0690091 838 bp mRNA linear EST 15-JUL-2002
 DEFINITION AGENCOURT_8064191 NIH_MGC_110 Homo sapiens CDNA clone IMAGE:6207831
 5', mRNA sequence.
 ACCESSION B0690091
 VERSION B0690091.1 GI:21815407
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 838)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLC2364 row: i column: 16
 High quality sequence stop: 670.

FEATURES
 source

1..838
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6207831"
 /tissue_type="NIH_MGC_110"
 /lab_host="DH10B (phage-resistant)"
 /notes="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; CDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."
 216 a -205 c 196 g 220 t 1 others

BASE COUNT
 ORIGIN

Query Match 49.2%; Score 765.8; DB 14; Length 838;
 Best Local Similarity 99.5%; Pred. No. 2.4e-228;
 Matches 767; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 531 TTGAGCGCTATCTACTGATGTAGTACGCGGTACAAAGTGAGCAGCATCACCCCTCACCAGC 590
 Db 1 TTGAGCGCTATCTACTGATGTAGTACGCGGTACAAAGTGAGCAGCATCACCCCTCACCAGC 60
 QY 591 AACTCCTACCTGATCTCTGTTCCAAAGTGGCAAGGAGGCAATCGCGGCGCCACAGATTG 650
 Db 61 AACTCCTACCTGATCTCTGTTCCAAAGTGGCAAGGAGGCAATCGCGGCGCCACAGATTG 120
 QY 651 ACAAGAAAGGAGCGGCTGCTCATGACCTTCTCTGAGGAGATGTGATCCGAGAATTTA 710
 Db 121 ACAAGAAAGGAGCGGCTGCTCATGACCTTCTCTGAGGAGATGTGATCCGAGAATTTA 180
 QY 711 ACTTAATGAGCTATACCAAGCGGCGCAAGAACTATCAAGGCTGGAGACAATATCCCTG 770
 Db 181 ACTTAATGAGCTATACCAAGCGGCGCAAGAACTATCAAGGCTGGAGACAATATCCCTG 240
 QY 771 AGGAGCAGCCTGTGCTTCAACCCCAACACAGTGTGATGGGAAACAGAGGATA 830
 Db 241 AGGAGCAGCCTGTGCTTCAACCCCAACACAGTGTGATGGGAAACAGAGGATA 300
 QY 831 AATAAGATCCTCAGCTTTGGCAGTGTCTCTCTCTCTCAATTTCCAGGCTTTTCCATAAC 890
 Db 301 AATAAGATCCTCAGCTTTGGCAGTGTCTCTCTCTCTCAATTTCCAGGCTTTTCCATAAC 360
 QY 891 CACAAGCCTGAGGCTGCAGCCTTTTATTTATGCTTTTCCCTTTGGCTGTGACTGGGTGGG 950
 Db 361 CACAAGCCTGAGGCTGCAGCCTTTTATTTATGCTTTTCCCTTTGGCTGTGACTGGGTGGG 420
 QY 951 CAGCATGAGCTTCTGATTTTAAAGAGGCAATCTAGGGAATTGTACAGCACCTTACAGGAA 1010
 Db 421 CAGCATGAGCTTCTGATTTTAAAGAGGCAATCTAGGGAATTGTACAGCACCTTACAGGAA 480
 QY 1011 GGCCTGCCATGCTGTGGCCCACTGTTTCACTGGAGCAAGAAAGATCTCATAGGACGA 1070

```
Db 481 GGCCTGCCATGCTGTGGCCCACTGTTTCACTGGAGCAAGAAAGAGATCTCATAGGACGGA 540
QY 1071 GGGGGAATGGTTTCCCTCCCAAGCTTGGGTCACTGTGTTAACTGCTTATCAGTATTCAG 1130
Db 541 GGGGGAATGGTTTCCCTCCCAAGCTTGGGTCACTGTGTTAACTGCTTATCAGTATTCAG 600
QY 1131 ACATCCCATGTTTCCATGAACCTCTGTGTTTTCATCATCTTCTTCTAGTTGACCTG 1190
Db 601 ACATCCCATGTTTCCATGAACCTCTGTGTTTTCATCATCTTCTTCTAGTTGACCTG 660
QY 1191 CACAGCTTGGTTAGACTAGATTAAACCTTAAGCTAAGTAACTGCTGGGTATAGAACGCTAA 1250
Db 661 CACAGCTTGGTTAGACTAGATTAAACCTTAAGCTAAGTAACTGCTGGGTATAGAACGCTAA 720
QY 1251 GAATTTTCCCAAGGACTCTTGTCTTCCCTTAAGCCCTTCGCTTCGTTT 1300
Db 721 GAATTTTCCCAAGGACTCTTGTCTTCCCTTAAGCCCTTCGCTTCGTTT 770

RESULT 12
BQ962329
-~C~S
INITIATION BQ962329 835 bp mRNA linear EST 21-AUG-2002
AGENCOURT 8820581 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:6379790
5', mRNA sequence.
ACCESSION BQ962329
VERSION BQ962329.1 GI:22377807
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 835)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LfCM2567 row: f column: 15
High quality sequence stop: 712.
Location/Qualifiers
1. .835
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6379790"
/clone_lib="NIH_MGC_42"
/tissue_type="epitheloid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"
BASE COUNT 228 a 192 c 210 g 205 t
ORIGIN
Query Match 49.1%; Score 763.2; DB 14; Length 835;
Best Local Similarity 97.7%; Pred. No. 1.5e-227;
Matches 774; Conservative 18; Mismatches 22; Indels 0; Gaps 0;
QY 265 GATGTTTCTCAGTCCATGTGATGATGAGAACCGCAGATCCATCTCTCTGATGACGTG 324
|||||
```

```
Db 1 GATGTTTCTCAGTCCATGTGATGATGAAGAACCGCAGATCCATCTCTCTGATGACGTG 60
QY 325 CAAACCCCCCTATATATATGGCCCTCAGTATATCAAGTACTTCAATGATAAAACCATTTGA 384
Db 61 CAAACCCCCCTATATATATGGCCCTCAGTATATCAAGTACTTCAATGATAAAACCATTTGA 120
QY 385 TGAGGAACATAGAACGGGACAAAGAGGTCACCTGGATTGTGGAGTTCTTTCCCAATTTGGTC 444
Db 121 TGAGGAACATAGAACGGGACAAAGAGGTCACCTGGATTGTGGAGTTCTTTCCCAATTTGGTC 180
QY 445 TAATGACTGCAATCATATTTGCCCTCTATCTATGCTGACCTCTCCCTTAAATACAACTGTAC 504
Db 181 TAATGACTGCAATCATATTTGCCCTCTATCTATGCTGACCTCTCCCTTAAATACAACTGTAC 240
QY 505 AGGCTAAATTTTGGGAAGTGGATGTTGGACGCTATATGATGTTAGTAGCGCGGTACAA 564
Db 241 AGGCTAAATTTTGGGAAGTGGATGTTGGACGCTATATGATGTTAGTAGCGCGGTACAA 300
QY 565 AGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCTGATCTCTGTTTCCAAGTGGCAA 624
Db 301 AGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCTGATCTCTGTTTCCAAGTGGCAA 360
QY 625 GGAGGCAATCGCGGGCCACAGATTGACAAAGAGGACGGGCTGTCTCATGGACCTTCTC 684
Db 361 GGAGGCAATCGCGGGCCACAGATTGACAAAGAGGACGGGCTGTCTCATGGACCTTCTC 420
QY 685 TGAGGAGAATGTGATCGGGAATTTAACTTAAATGAGCTATACCAAGGGGCGCAAGAACT 744
Db 421 TGAGGAGAATGTGATCGGGAATTTAACTTAAATGAGCTATACCAAGGGGCGCAAGAACT 480
QY 745 ATCAAGGCTGGGACAAATATCCCTGAGGAGACGCTGTGGCTTCAACCCCCACACAGT 804
Db 481 ATCAAGGCTGGGACAAATATCCCTGAGGAGACGCTGTGGCTTCAACCCCCACACAGT 540
QY 805 GTCAGATGGGGAACAAAGAGGATAAATAAGATCTCTCACTTTGGCAGTGTCTCTCTCC 864
Db 541 GTCAGATGGGGAACAAAGAGGATAAATAAGATCTCTCACTTTGGCAGTGTCTCTCTCC 600
QY 865 TGTCATTTCCAGGCTCTTCCATTAACCAACCAAGCTGAGGCTGAGGCTTTTATTTATGTT 924
Db 601 TGTCATTTCCAGGCTCTTCCATTAACCAACCAAGCTGAGGCTGAGGCTTTTATTTATGTT 660
QY 925 TTCCCTTTGGCTGTGACTGGGTGGGCGAGCATGTCAGCTTCTGATTTTAAAGAGCATCTA 984
Db 661 TTCCCTTTGGCTGTGACTGGGTGGGCGAGCATGTCAGCTTCTGATTTTAAAGAGCATCTA 720
QY 985 GGAATTTGTCAGGACCTTACAGGAAGGCTGCCATCTGTGGCCAACTGTTTCACTGGA 1044
Db 721 GGAATTTGTCAGGACCTTACAGGAAGGCTGCCATCTGTGGCCAACTGTTTCACTG 780
QY 1045 GCAAGAAAGAGA 1056
Db 781 GGAGCAAGGAAA 792

RESULT 13
BQ920164
LOCUS BQ920164 913 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT_8918912 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383675
5', mRNA sequence.
ACCESSION BQ920164
VERSION BQ920164.1 GI:22334862
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 913)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
```

Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM2577 row: h column: 12
 High quality sequence stop: 684.
 Location/Qualifiers
 1. .913
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6383675"
 /clone_lib="NIH_MGC_40"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dt priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 234 a 241 c 215 g 221 t 2 others

FEATURES
source

BASE COUNT 234 a 241 c 215 g 221 t 2 others
 ORIGIN
 Query Match 48.3%; Score 751.6; DB 14; Length 913;
 Best Local Similarity 92.7%; Pred. No. 6.9e-224;
 Matches 844; Conservative 0; Mismatches 31; Indels 35; Gaps 4;
 QY 51 CCAGAAAGATGGCGGCTTGGACCTCTAATTCCTGCTGTTATTCGGTCCGGGACCTTT 110
 DB 1 CCAGAAAGATGGCGGCTTGGACCTCTAATTCCTGCTGTTATTCGGTCCGGGACCTTT 60
 QY 111 CAGATGGCTGGCCCAACCTTACTACTTCTGTCGGCCCTGCTCTCTGCTGCTTCCTAC 170
 DB 61 CAGATGGCTGGCCCAACCTTACTACTTCTGTCGGCCCTGCTCTCTGCTGCTTCCTAC 120
 QY 171 TCGTGAGGAACCTGCGCGGCTCTGCGACGGTCTGCCACCCACCCAGCGGAGAGCGT 226
 DB 121 TCGTGAGGAACCTGCGCGGCTCTGCGACGGTCTGCCACCCACCCAGCGGAGAGCGT 180
 C 227 -----AACCGTGTGACTTTGACTGGAGAGAGTGAGA 260
 181 CAGATGCACCCCATGCTGGGAACGCAACCCCTTTGACAGAAAGAGAGAGTGAGA 240
 261 TCCTGATGTTCTCAGTGCCATTGTGATGATGAAGAACCCGACATCCATGTTCTGATGA 320
 241 TCCTGATGTTCTCAGTGCCATTGTGATGATGAAGAACCCGACATCCATGTTCTGATGA 300
 321 CGTGCAACCCCCCTATATATATGGGCGCTGAGTATATCAAGTACTTCAATGATAAACCA 380
 301 CGTGCAACCCCCCTATATATATGGGCGCTGAGTATATCAAGTACTTCAATGATAAACCA 360
 381 TTGATGAGGAACCTAGAACGGGCAAGAGGTCTACTTGGAGTTGGAGTTCTTTGCCAATT 440
 361 TTGATGAGGAACCTAGAACGGGCAAGAGGTCTACTTGGAGTTGGAGTTCTTTGCCAATT 420
 441 GGTCTAATGACTGCCAATCATTTGCCCTTATCTATGCTGACCTCTCCCTTAAATACAAC 500
 421 GGTCTAATGACTGCCAATCATTTGCCCTTATCTATGCTGACCTCTCCCTTAAATACAAC 480
 501 GTACAGGCGTAATTTTGGGAAGGTGATGTTGGAGCGCTATACCTGATGTAGTACGCGGT 560
 481 GTACAGGCGTAATTTTGGGAAGGTGATGTTGGAGCGCTATACCTGATGTAGTACGCGGT 540
 561 ACAAGTGAACATCAACCCCTCACCAGCAACTCCCTACCTGATGCTCTGTTCCAGAGGTG 620
 541 ACAAGTGAACATCAACCCCTCACCAGCAACTCCCTACCTGATGCTCTGTTCCAGAGGTG 600
 621 GCAAGGAGCAATGCGGGGCCACAGATTGACAAGAAAGACGGGCTGCTCATGACCT 680

Db 601 GCAAGGAGCAATGCGGGGCCACAGATTGACAAGAGGCGGCTGTCTCATGACCT 660
 QY 681 TCTCTGAGGAGAAATGTCGCGAGAAATTTAACTAAATGAGCTATACACGCGGCAAGA 740
 Db 661 TCTCTGAGGAGAAATGTCGCGAGAAATTTAACTAAATGAGCTATACACGCGGCAAGA 720
 QY 741 AACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACA 800
 Db 721 AACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACA 780
 QY 801 CAGTGTCAAGTGGGAAAACAAGGATTAATTAAGATCCCTCACTTTGGCAGTCTCTCC 859
 Db 781 CAGTGTCAAGTGGGAAAACAAGGATTAATTAAGATCCCTCACTTTGGCAGGCTTC 840
 QY 860 TCT---CCTGCAATTCAGGCTCTTCCATAACCAACA--GCCTGAGGCTGCGAGCTTTT 915
 Db 841 CTTCTCCCTGTCATTCAGGCTCTTCCATAACCAACAAGGCTGAGGCTGCGAGNCTTT 900
 QY 916 ATTTATGTTT 925
 Db 901 TTTTATTAT 910

RESULT 14

LOCUS BQ687916 894 bp mRNA linear EST 15-JUL-2002
 DEFINITION AGENCOURT_8344073 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6248661
 5', mRNA sequence.
 ACCESSION BQ687916
 VERSION BQ687916.1 GI:21813232
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 894)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM2388 row: n column: 22
 High quality sequence stop: 645.
 Location/Qualifiers
 1. .894
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6248661"
 /clone_lib="NIH_MGC_110"
 /tissue_type="ductal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dt priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

FEATURES

source

BASE COUNT 236 a 204 c 222 g 231 t 1 others
 ORIGIN
 Query Match 48.3%; Score 750.8; DB 14; Length 894;
 Best Local Similarity 99.1%; Pred. No. 1.2e-223;
 Matches 807; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

QY	369	ATGATAAAACCATGATGAGGAACATAGAACGGGACAGAGGGTCACTTGGATTGGAGT	428
Db	1	ATGATAAAACCATGATGAGGAACATAGAACGGGACAGAGGGTCACTTGGATTGGAGT	60
QY	429	TCTTTGCCAATTTGGTCTAATGACTGCCAATATTTGGCCCTATCTATGCTGACCTCTCC	488
Db	61	TCTTTGCCAATTTGGTCTAATGACTGCCAATATTTGGCCCTATCTATGCTGACCTCTCC	120
QY	489	TTAATAACAACGTACAGGGCTAAATTTGGGAAGTGGATGTTGACGCTATACCTGATG	548
Db	121	TTAATAACAACGTACAGGGCTAAATTTGGGAAGTGGATGTTGACGCTATACCTGATG	180
QY	549	TTAGTACCGGTACAAAAGTACACATCACCTCCACCAAGCACTCCCTACCTGATCC	508
Db	181	TTAGTACCGGTACAAAAGTACACATCACCTCCACCAAGCACTCCCTACCTGATCC	240
QY	609	TGTTCCRAAGTGGCAAGGAGCAATGCGGCGCCACAGATTCACAGAAGACGCGGTG	668
Db	241	TGTTCCRAAGTGGCAAGGAGCAATGCGGCGCCACAGATTCACAGAAGACGCGGTG	300
QY	669	TCTCATGACCTCTCTCTGAGGAGATGTGATCCGAGAAATTTAACTTAAATGAGCTATAC	728
Db	301	TCTCATGACCTCTCTCTGAGGAGATGTGATCCGAGAAATTTAACTTAAATGAGCTATAC	360
QY	729	AGCGGGCCAGAAACTATCAAGGCTGGAGACAATATCCCTGAGGAGAGCCTGTGGCTT	788
Db	361	AGCGGGCCAGAAACTATCAAGGCTGGAGACAATATCCCTGAGGAGAGCCTGTGGCTT	420
QY	789	CAACCCCCACACACGTGTCAGATGGGGAACCAAGAGGATAAATAAGATCCTCACTTGG	848
Db	421	CAACCCCCACACACGTGTCAGATGGGGAACCAAGAGGATAAATAAGATCCTCACTTGG	480
QY	849	GCAGTGTCTCTCTCTGTCATTCAGGCTCTTTCCATAACCAAGCCTGAGGCTGCA	908
Db	481	GCAGTGTCTCTCTCTGTCATTCAGGCTCTTTCCATAACCAAGCCTGAGGCTGCA	540
QY	909	GCCTTTTATTTATGTTTCCCTTTGGCTGTGACTGGTGGGCGAGCATCGCTTCTGAT	968
Db	541	GCCTTTTATTTATGTTTCCCTTTGGCTGTGACTGGTGGGCGAGCATCGCTTCTGAT	600
QY	969	TTTAAAGAGGATCTAGGGAATTTGTCAGGACCCCTACAGGAAGGCTGCATGCTGTGG	1028
Db	601	TTTAAAGAGGATCTAGGGAATTTGTCAGGACCCCTACAGGAAGGCTGCATGCTGTGG	660
QY	1029	CAACTGTTTCACTGGAGCAAGAGATCTCATAGGACGGA-GGGGGAATGTTTCCC	1087
Db	661	CAACTGTTTCACTGGAGCAAGAGATCTCATAGGACGGAAGGGAATGTTTCCC	720
QY	1088	TCCAAAGCTT-GGGTCAGTGTGTTAACT-GCTTATCAGCTATTACAGACATCTCCATGGTTT	1145
Db	721	TCCAGGCTTGGGGTCAGTGTGTTAACTGCTTATCAGCTATTACAAACATCTCCATGGTTT	780
QY	1146	CTCCATGAA-ACTCTGT-GGTTTCATCTCCCTT	1177
Db	781	CTCCATGAANACTCTGTGGGTTTCATCACTCTT	814
RESULT	15		
LOCUS	AL569941/c		
DEFINITION	AL569941 LTI_NFL006_PL2 Homo sapiens cdna clone CS0DI003YH21 3	786 bp	linear EST 16-FEB-2001
ACCESSION	AL569941		prime, mRNA sequence.
VERSION	AL569941.1	GI:12925779	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
	1 (bases 1 to 786)		
	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
FEATURES	Location/Qualifiers		
Source	1..786		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DI003YH21"		
	/clone_lib="LTI_NFL006_PL2"		
	/tissue_type="placenta"		
	/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA		
	was primed with a NotI-oligo(dT) primer. Five prime end		
	enriched, double-stranded cDNA was digested with Not I and		
	cloned into the Not I and Eco RV sites of the pCMVSPORT 6		
	vector. Library was normalized. Library was constructed by		
	Life Technologies. Contact : Feng Liang Life Technologies,		
	a division of Invitrogen 9800 Medical Center Drive,		
	Rockville, Maryland 20850, USA Fax : (1) 301 610 8371		
	Email : fliang@lifetech.com URL :		
	http://fulllength.invitrogen.com		
BASE COUNT	226 a 181 c 171 g 204 t		
ORIGIN			
Query Match	48.0%; Score 746.2; DB 9; Length 786;		
Best Local Similarity	98.9%; Pred. No. 3.2e-222;		
Matches	778; Conservative 4; Mismatches 2; Indels 3; Gaps 3;		
QY	677	ACCTCTCTGAGGAGAAATGTGATCCGAGAAATTTAACTTAAATGAGCTATACCAAGCGGCC	736
Db	786	MCCTTTCTGAGGAGAAATGTGATCCGAGAAATTTAACTTAAATGAGCTATACCAAGCGGCC	727
QY	737	AAGAAACTATCAAAAGGCTGGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCC	796
Db	726	AAGAAACTATMAAAGGCTGGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCC	667
QY	797	ACCACAGTCTCAGATGGGAAAACAAGAGGATAAATAAGATCCTCACATTTGGCAGTGT	856
Db	666	ACCACAGTCTCAGATGGGAAAACAAGAGGATAAATAAGATCCTCACATTTGGCAGTGT	607
QY	857	TCTCTCTCTCAATTCAGGCTCTTTCCATAACCAAGCCTGAGGCTGCAGCCTTTTAA	916
Db	606	TCTCTCTCTCAATTCAGGCTCTTTCCATAACCAAGCCTGAGGCTGCAGCCTTTTAA	547
QY	917	TTATGTTTTTCCCTTTGGCTGTGACTGGTGGGCGAGCATGACGCTTCTGATTTTAAAGA	976
Db	546	TTATGTTTTTCCCTTTGGCTGTGACTGGTGGGCGAGCATGACGCTTCTGATTTTAAAGA	487
QY	977	GGCATCTAGGGAATGTCAGGCACCTCAGGAAGGCTGCCATGCTGTGGCCAACTGTT	1036
Db	486	GGCATCTAGGGAATGTCAGGCACCTCAGGAAGGCTGCCATGCTGTGGCCAACTGTT	427
QY	1037	TCAGTGGACAGAAAGAGATCTATGAGCAGGAGGGGAAATGTTTCCCTCCAAGCTT	1096
Db	426	TCAGTGGACAGAAAGAGATCTATGAGCAGG- GGGGAAATGTTTCCCTCCAAGCTT	368
QY	1097	GGGTCACTGTTTAACTGCTTATCAGCATCTCCATGTTTCTCCATGAAC	1156
Db	367	GGGTCACTGTTTAACTGCTTATCAGCATCTCCATGTTTCTCCATGAAC	308
QY	1157	TCGTGGTTTTCATCTCTTCTTACCTGACGACAGCTTGGTTAGACCTAGATTTAA	1216
Db	307	TCGTGGTTTTCATCTCTTCTTACCTGACGACAGCTTGGTTAGACCTAGATTTAA	248
QY	1217	CCCTAAGGTAAGATGCTGGGGTATACACGCTAAAGAAATTTCCCCCAAGGACTCTTGCCT	1276
Db	247	CCCTAAGGTAAGATGCTGGGGTATACACGCTAAAGAAATTTCCCCCAAGGACTCTTGCCT	188
QY	1277	CCTTAAGCCCTCTCTGGCTTTCGT-TTATGTTCTTATTAAAGATTAAGCTTAACCTTGTG	1335
Db	187	CCTTAAGCCCTCTCTGGCTTTCGT-TTATGTTCTTATTAAAGATTAAGCTTAACCTTGTG	128

```

Qy 1336 GCTAGTCCTAAGGAGAAACCTTTAACCAAGTTTATCATTTGAAGACAATATTGAAC 1395
Db 127 GCTAGTCCTAAGGAGAAACCTTTAACCAAGTTTATCATTTGAAGACAATATTGAAC 68
Qy 1396 AACCCCTAT-TTTGTGGGGATTGAGAGGGGTGAATAGAGGCTTGAGACTTTCCTTTGT 1454
Db 67 AACCCCTATTTGTGGGGATTGAGAGGGGTGAATAGAGGCTTGAGACTTTCCTTTGT 8
Qy 1455 GTGGTAG 1461
Db 7 GTGGTAG 1

```

Search completed: July 12, 2003, 11:06:13
Job time : 1612 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 08:35:42 ; Search time 2787 Seconds
(without alignments)
16237.841 Million cell updates/sec

Title: US-09-954-846-4
Perfect score: 1555
Sequence: 1 AGGGGAGGGGGCGGAGACC.....TAATAAAGATTGGGATTA 1555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1381.2	88.8	1644	9	AF059753	AF059753 Homo sapi
2	1379.4	88.7	1646	9	BC000666	BC000666 Homo sapi
3	1370.2	88.1	1631	6	AX136115	AX136115 Sequence
4	1370	88.1	1620	6	AX375964	AX375964 Sequence
5	1355.4	87.2	1669	9	AF132965	AF132965 Homo sapi
6	962.2	61.9	189271	9	AL138752	AL138752 Human DNA
7	934.2	60.1	112389	9	AL138752	AL138752 Human DNA
8	868	55.8	110116	9	AC073598	AC073598 Homo sapi
9	868	55.8	130642	2	AP001931	AP001931 Homo sapi
10	868	55.8	158349	2	AC009438	AC009438 Homo sapi
11	868	55.8	161397	2	AC090218	AC090218 Homo sapi
12	868	55.8	170540	2	AC092707	AC092707 Homo sapi
13	868	55.8	199992	2	AC021522	AC021522 Homo sapi
14	856	55.0	153394	2	AP000727	AP000727 Homo sapi
15	571.6	36.8	166341	2	AC013381	AC013381 Homo sapi
16	548.4	35.3	1978	10	BC013544	BC013544 Mus muscu
17	544.2	35.0	200653	2	AC127695	AC127695 Mus muscu
18	463	29.8	726	6	AX136467	AX136467 Sequence
19	328.8	21.1	395	11	GI4613	GI4613 human STS S
20	258.4	16.6	307	6	AX396397	AX396397 Sequence
21	252.2	16.2	5423	6	AX336311	AX336311 Sequence
22	252.2	16.2	5423	9	AB002382	AB002382 Human mrn
23	246	15.8	117711	9	AP000662	AP000662 Homo sapi
24	246	15.8	161397	2	AC090218	AC090218 Homo sapi
25	221.4	14.2	1730	9	BC008500	BC008500 Homo sapi
26	158.6	10.2	183656	2	AC111385	AC111385 Rattus no
27	156.2	10.0	149481	2	AC102678	AC102678 Mus muscu
28	156.2	10.0	216237	2	AC121786	AC121786 Mus muscu
29	153.4	9.9	81994	2	AC096003	AC096003 Rattus no
30	143.2	9.2	183656	2	AC111385	AC111385 Rattus no
31	127	8.2	127	11	G62115	G62115 1897 Human
32	110.2	7.1	153394	2	AP000727	AP000727 Homo sapi
33	95.8	6.2	1223	3	AY118488	AY118488 Drosophil
34	95.8	6.2	31857	2	AC020299	AC020299 Drosophil
35	95.8	6.2	45105	3	AC004311	AC004311 Drosophil
36	95.8	6.2	160106	3	AC099014	AC099014 Drosophil
37	95.8	6.2	173347	3	AC099013	AC099013 Drosophil
38	95.8	6.2	248096	3	AE003795	AE003795 Drosophil
39	89.6	5.8	81994	2	AC096003	AC096003 Rattus no
40	53.6	3.4	36882	3	U21324	U21324 Caenorhabdi
41	43.2	2.8	161888	9	HS70501	AL035251 Human DNA
42	42.6	2.7	14461	6	AX256438	AX256438 Sequence
43	42.4	2.7	5204	5	GGU58204	U58204 Gallus gall
44	41.8	2.7	67446	2	AC103818	AC103818 Homo sapi
45	41.8	2.7	184302	9	AL162393	AL162393 Human DNA

ALIGNMENTS

RESULT 1
AF059753
LOCUS AF059753 1644 bp mRNA linear PRI 02-JAN-2001
DEFINITION Homo sapiens clone 008a05 My009 protein mRNA, complete cds.
ACCESSION AF059753
VERSION AF059753.1 GI:12001957
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1644)
AUTHORS Mao,Y.M., Xie,Y., Zhou,Z.X., Ying,K. and Zheng,Z.H.
TITLE Direct Submission
JOURNAL Submitted (16-APR-1998) Institute of Genetics, School of Life

Science, Fudan University, 220 Handan Rd., Shanghai 200433,
P.R.China

FEATURES
source Location/Qualifiers

1..1644
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="008a05"
/tissue_type="brain"
/dev_stage="fetus"
28..1146
/codon_start=1
/product="My009 protein"
/protein_id="A043122.1"
/db_xref="GI:12001958"
/translation="MAVLAPLIALVSVPRISRWLAQPYLLSALLSNAFLVLRKLPP
LCHGLPQRENDPCDFDWEIVELIMFSAIVMMKNRSITVEQHGIFNMFVKVANT
ILFRLDIRMLGLITLGLVEILMTKPLPYMGPEYIKYFNDRKIDEEELKRVWIV
EFFANWSDCOSFAPYADLSLKYNTGLNFGKVDVGRVTVSTYKSTVSLTKQLP
TLILPOGSKEMRRPOISKKGRAVSWTFSEENVIREFNELNYORAKLSKAGDNIPE
EQPVASTPPOCMGKTRINKILITLAVLPLLSIPGSFHNKPEAAAFYLCFFPGCDWV
GQHAASDFKASRELSTGLQGLPCGGLFHWKKEIS"

BASE COUNT 412 a 401 c 385 g 446 t

Juery Match 88.8%; Score 1381.2; DB 9; Length 1644;
Best Local Similarity 92.8%; Pred. No. 0; Mismatches 3; Indels 115; Gaps 2;
Matches 1318; Conservative 0;

Qy 33 GAGCAGTGGCGGTTACGGCGAAAGATGGCGGTCTTGGCACCTCTAAATGCTCTCGTGT 92
Db 2 GAGCAGTGGCGGTTACGGCGAAAGATGGCGGTCTTGGCACCTCTAAATGCTCTCGTGT 61
Qy 93 ATTCGGTGGCGGATTTACGATGGCTCGGCCAACCTTACTACCTTCTGTGCGGCCGTC 152
Db 62 ATTCGGTGGCGGATTTACGATGGCTCGGCCAACCTTACTACCTTCTGTGCGGCCGTC 121
Qy 153 TCTCTGCTGCTTCTACTCGTGGAGAACTCGCGCGCTCTGCCACGCTCTGCCACCC 212
Db 122 TCTCTGCTGCTTCTACTCGTGGAGAACTCGCGCGCTCTGCCACGCTCTGCCACCC 181
Qy 213 AACGCGAAGCGGTAAACCGGTGACTTTGACTGGAGAGAAGTGGAGATCTCTGATGTTTC 272
Db 182 AACGCGAAGCGGTAAACCGGTGACTTTGACTGGAGAGAAGTGGAGATCTCTGATGTTTC 241
Qy 273 TCAGTGCCATTGTGATGATGAAGAACCGCAGATCA ----- 308
Db 242 TCAGTGCCATTGTGATGATGAAGAACCGCAGATCACTACTGAGGCAACATATAGGCA 301
Qy 309 ----- 308
Db 302 ACATTTTCATGTTAGTAAAGTGGCGCAACACAAATCTTTTCTCCGCTTGGATATTCGCA 361
Qy 309 -----TCTTCTGATGACGTGCAACCCCGCTAT 338
Db 362 TGGGCTACTTTACATCACACTCTGCATAGTCTTCTGATGACGTGCAACCCCGCTAT 421
Qy 339 ATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAAACCATTGATGAGGAACATGAAC 398
Db 422 ATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAAACCATTGATGAGGAACATGAAC 481
Qy 399 GGGACAAGGGTCACTTGGATTTGGAGTTCTTTGCCAATTTGGCTAATGACTGCCAAT 458
Db 482 GGGACAAGGGTCACTTGGATTTGGAGTTCTTTGCCAATTTGGCTAATGACTGCCAAT 541
Qy 459 CATTTGCCCTTATCTATGCTGACCTCTCCCTTAAATACAACTGTACAGGGCTAAATTTTG 518
Db 542 CATTTGCCCTTATCTATGCTGACCTCTCCCTTAAATACAACTGTACAGGGCTAAATTTTG 601
Qy 519 GGAAGGTGGATTTGGAGCGCTATCTATGATGTTAGTACGGCGGTACAAAGTGAACATCAC 578
Db 602 GGAAGGTGGATTTGGAGCGCTATCTATGATGTTAGTACGGCGGTACAAAGTGAACATCAC 661
Qy 579 CCCTACCAAGCAACTCCCTACCTCTGATCTCTTCCAGGTGGCAAGGAGCAATCGGC 638

RESULT 2
BC000666
LOCUS
DEFINITION

BC000666 1646 bp mRNA linear PRI 12-JUL-2001
Homo sapiens, CGI-31 protein, clone MGC:713 IMAGE:3349733, mRNA,
complete cds.

Db 662 CCCTCACCAGCAACTCCCTACCTGATCTGTTCACAGGTGCACAGGAGCAATCGGC 721
Qy 639 GGGCCAGATTTGACAAAGAAAGGACGGCTGTCTCATGGACCTTCTCTGAGGAGAATGTGA 698
Db 722 GGGCCAGATTTGACAAAGAAAGGACGGCTGTCTCATGGACCTTCTCTGAGGAGAATGTGA 781
Qy 699 TCCGAGAAATTTAACTTAATGAGCTATACCAGGGGGCCCAAGAACTATCAAGAGCTGGAG 758
Db 782 TCCGAGAAATTTAACTTAATGAGCTATACCAGGGGGCCCAAGAACTATCAAGAGCTGGAG 841
Qy 759 ACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCGCCACACAGTGTTCAGATGGGAAA 818
Db 842 ACAATATCCCTGAGGAGCAGCTGTGGCTTCAA -CCCCACCAGTGTTCAGATGGGAAA 900
Qy 819 ACAAGAGGATTAATAAGATCTCTACTTTGGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCT 878
Db 901 ACAAGAGGATTAATAAGATCTCTACTTTGGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
Qy 879 TCTTTCCATTAACCAACAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 938
Db 961 TCTTTCCATTAACCAACAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1020
Qy 939 GACTGGTGGGCGCAGCATGCAGCTTCTGATTTTAAAGAGGACATCTAGGGAATGTCTACGC 998
Db 1021 GACTGGTGGGCGCAGCATGCAGCTTCTGATTTTAAAGAGGACATCTAGGGAATGTCTACGC 1080
Qy 999 ACCCTACAGAAAGGCGCTGCTGCTGCGCAACTGTTTCACTGGAGCAAGAAAGAGATC 1058
Db 1081 ACCCTACAGAAAGGCGCTGCTGCTGCGCAACTGTTTCACTGGAGCAAGAAAGAGATC 1140
Qy 1059 TCATAGGAGGAGGGGGAATGGTTTCCCTCCCAAGCTTGGGTCAGTGTGTAACTGCTTA 1118
Db 1141 TCATAGGAGGAGGGGGAATGGTTTCCCTCCCAAGCTTGGGTCAGTGTGTAACTGCTTA 1200
Qy 1119 TCAGCTATTTCAGACATCTCCATGTTTCCCATGAAACTCTGTGGTTTTCATATCTCTTC 1178
Db 1201 TCAGCTATTTCAGACATCTCCATGTTTCCCATGAAACTCTGTGGTTTTCATATCTCTTC 1260
Qy 1179 TTAGTTGACCTGCACAGCTTGGTTAGACCTAGATTTTAAACCTTAAGGTAAGATGCTGGGT 1238
Db 1261 TTAGTTGACCTGCACAGCTTGGTTAGACCTAGATTTTAAACCTTAAGGTAAGATGCTGGGT 1320
Qy 1239 ATAGAACGCTTAAGAAATTTCCCAAGGACTCTTGTCTTCCCTTAAGCCCTTCTGGCTCGT 1298
Db 1321 ATAGAACGCTTAAGAAATTTCCCAAGGACTCTTGTCTTCCCTTAAGCCCTTCTGGCTTCGT 1380
Qy 1299 TTATGCTCTTCATTAAGATTAAGCCTTAACCTTTGTGCTAGTCTCTTAAGGAGAAACCTTT 1358
Db 1381 TTATGCTCTTCATTAAGATTAAGCCTTAACCTTTGTGCTAGTCTCTTAAGGAGAAACCTTT 1440
Qy 1359 AACCAAAAGTTTTCATTAAGCAATAATTAAGCAACCCCTATTTTGTGGGATG 1418
Db 1441 AACCAAAAGTTTTCATTAAGCAATAATTAAGCAACCCCTATTTTGTGGGATG 1500
Qy 1419 AGAAGGGTGAATAGAGCTTGAAGCTTTCCCTTTGTGTGTGAGGACTTGGAGGAGAAATC 1478
Db 1501 AGAAGGGTGAATAGAGCTTGAAGCTTTCCCTTTGTGTGTGAGGACTTGGAGGAGAAATC 1560
Qy 1479 CCCTGACCTTTCACCTAACCTCTGACATCTCCCAACACCCAGTGTGATGGCTTTCCGTAA 1538
Db 1561 CCCTGACCTTTCACCTAACCTCTGACATCTCCCAACACCCAGTGTGATGGCTTTCCGTAA 1620
Qy 1539 TAAAAAGATTGGGATT 1554
Db 1621 TAAAAAGATTGGGATT 1636

ACCESSION	BC000666	QY	170	CTCGTGAGAAACGCGCCGCTCTGCGACAGTCTGCCACACCCACCGAAGACGGTAC	229
VERSION	MG000666.1	Db	121	CTCGTGAGAAACGCGCCGCTCTGCGACAGTCTGCCACACCCACCGAAGACGGTAC	180
KEYWORDS	MGC.	QY	230	CCGTGTGACTTTGACTGGAGAGAGTGGAGATCCTGATGTTTCTCAGTGCCCATTTGATG	289
SOURCE	Homo sapiens.	Db	181	CCGTGTGACTTTGACTGGAGAGAGTGGAGATCCTGATGTTTCTCAGTGCCCATTTGATG	240
ORGANISM	Homo sapiens	QY	290	ATGAAGAACCAGATGCA	308
REFERENCE	Submitted (15-NOV-2000)	Db	241	ATGAAGAACCAGATGCA	300
AUTHORS	Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	QY	309	-----TGTTCCTGATGACGTGCAAAACCCCTTATATATATGGCCCTCAGTAT	355
TITLE	NIH-MGC Project URL: http://mgi.nci.nih.gov	Db	301	AAAGTGGCAACAAATTTCTTCGCTTGGATATTCGCATGGGCTACTTTACATC	360
JOURNAL	Contact: MGC help desk	QY	309	-----TGTTCCTGATGACGTGCAAAACCCCTTATATATATGGCCCTCAGTAT	355
	Email: gcapbs@mail.nih.gov	Db	361	ACACTCTGCATAGTGTCTCTGATGACGTGCAAAACCCCTTATATATATGGCCCTCAGTAT	420
	Tissue Procurement: ATCC	QY	356	ATCAAGTACTTCAATGATAAAACCATTTGATGAGGAACCTAGAACGGGACAGGGTCACT	415
	cDNA Library Preparation: Rubin Laboratory	Db	421	ATCAAGTACTTCAATGATAAAACCATTTGATGAGGAACCTAGAACGGGACAGGGTCACT	480
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	QY	416	TGGATTCTGGAGTTCCTTGGCAATTTGGTCTTAATGACTGCCAATCATTTGCCCTCTATCTAT	475
	DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;	Db	481	TGGATTCTGGAGTTCCTTGGCAATTTGGTCTTAATGACTGCCAATCATTTGCCCTCTATCTAT	540
	Web site: http://www.nisc.nih.gov/	QY	476	GCTGACCTCTCCCTTAAATACAACTGTACAGGGCTAAATTTGGGAAGTGGATGTCGA	535
	Contact: nisc.mgc@nih.gov	Db	541	GCTGACCTCTCCCTTAAATACAACTGTACAGGGCTAAATTTGGGAAGTGGATGTCGA	600
	Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, J.C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.	QY	536	CGCTATCTGATGTAGTACGCGGTACAAAGTGGAGACATCACCCCTCACCAAGCACTC	595
		Db	601	CGCTATCTGATGTAGTACGCGGTACAAAGTGGAGACATCACCCCTCACCAAGCACTC	660
		QY	596	CCTACCTCTGATCTCTTCCAAAGTGGCAAGAGGCAATTCGCGGGGCCACAGATTGACAAG	655
		Db	661	CCTACCTCTGATCTCTTCCAAAGTGGCAAGAGGCAATTCGCGGGGCCACAGATTGACAAG	720
		QY	656	AAAGGACGGGTCTCTCATGGACCTCTCTGAGGAGATGTATCCGAGAAATTTAACTTA	715
		Db	721	AAAGGACGGGTCTCTCATGGACCTCTCTGAGGAGATGTATCCGAGAAATTTAACTTA	780
		QY	716	AATGAGCTATACCAAGCGGGCCAAAGAACTATCAAGGCTGGAGACATATCCCTGAGGAG	775
		Db	781	AATGAGCTATACCAAGCGGGCCAAAGAACTATCAAGGCTGGAGACATATCCCTGAGGAG	840
		QY	776	CAGCCTGTGGCTTCAACCCCAACAGTGTACAGTGGGAAACACAGAGGATAAATAA	835
		Db	841	CAGCCTGTGGCTTCAACCCCAACAGTGTACAGTGGGAAACACAGAGGATAAATAA	900
		QY	836	GATTCCTCACTTTGGCAGTGTCTCTCTCTCAATTCAGGCTCTTCCATATACCAACAA	895
		Db	901	GATTCCTCACTTTGGCAGTGTCTCTCTCTCAATTCAGGCTCTTCCATATACCAACAA	960
		QY	896	GCTGAGGCTGCGACCTTTTATTTATGTTTTTCCCTTTGGCTGTGACTGGGTGGGCGACA	955
		Db	961	GCTGAGGCTGCGACCTTTTATTTATGTTTTTCCCTTTGGCTGTGACTGGGTGGGCGACA	1020
		QY	956	TGCAGCTTCTGATTTTAAAGAGGATCTAGGAAATTCAGGCAACCTACAGGAAGGCT	1015
		Db	1021	TGCAGCTTCTGATTTTAAAGAGGATCTAGGAAATTCAGGCAACCTACAGGAAGGCT	1080
		QY	1016	GCCATCTGTGGCAACTGTTTTCACCTGGAGCAAGAGAGATCTCATAGGACGAGGGG	1075
		Db	1081	GCCATCTGTGGCAACTGTTTTCACCTGGAGCAAGAGAGATCTCATAGGACGAGGGG	1140
		QY	1076	AAATGTTTTCCCTCCCAAGCTTGGGTGAGTGTGTTAACTGCTTATCAGTATTCAGACATC	1135
		Db	1141	AAATGTTTTCCCTCCCAAGCTTGGGTGAGTGTGTTAACTGCTTATCAGTATTCAGACATC	1200

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 5 Row: d Column: 22
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7705725.

FEATURES

Location/Qualifiers
 1..1646
 /organism="Homo sapiens"
 /db_xref="LocustID:51075"
 /db_xref="taxon:9606"
 /clone="MGC:713 IMAGE:3349733"
 /tissue.type="Eye, retinoblastoma"
 /clone.lib="NIR_MGC16"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 10..900
 /codon_start=1
 /product="CGI-31 protein"
 /protein_id="AAH00666.1"
 /db_xref="GI:12653761"
 /translation="MAVLAPLIALVYSPRLSRWLAPYLLSALLSAAFLVLRKLPP
 LCHGLPTQREDNCPDFWREVEILFLSAIVMKNRISITVEQHTGNFMFSKVANT
 ILFLPDLIRMSGLYTLICLVFTMKPLMGPEYIKYFNDKTIDELERDRKVTWIV
 EFFANWSDCSFAPYLDLSLKYNCTGLNFKGVQVGYDYSTRYKYSTPLTKOLP
 TLILFOGCKEARRRQIDKKGRVSWTFSEENVIREFNELNYQRKKLSKAGDNIPE
 EQPVASTTIVSDGKNNKDK"
 429 a 395 c 376 g 446 t

CDS

BASE COUNT 429 a 395 c 376 g 446 t
 ORIGIN
 Query Match 88.7%; Score 1379.4; DB 9; Length 1646;
 Best Local Similarity 92.9%; Score No. 0;
 Matches 1504; Conservative 0; Mismatches 1; Indels 114; Gaps 1;

QY 50..GCCGAAAGATGGCGCTCTTGGCACCTCTAAATGCTCTCGTGTATTCGGTGGCGGACATT 109
 Db 1 GCCGAAAGATGGCGCTCTTGGCACCTCTAAATGCTCTCGTGTATTCGGTGGCGGACATT 60
 QY 110 TCAGGATGGCTGGCCCAACCTTACTACCTTCTGTCGGCCCTCTCTCTGCTGCTTCCCTA 169
 Db 61 TCAGGATGGCTGGCCCAACCTTACTACCTTCTGTCGGCCCTCTCTCTGCTGCTTCCCTA 120

QY 1136 TCCATGGTTTCTCCATGAACCTCTGTGGTTTCATCATCTCTTCTAGTTGACCTGCACAG 1195
Db 1201 TCCATGGTTTCTCCATGAACCTCTGTGGTTTCATCATCTCTTCTAGTTGACCTGCACAG 1260
QY 1196 CTGGTTAGACCTAGATTTAAACCCCTAAGTAAGATGCTGGGTATAGAACGCTAAGAATT 1255
Db 1261 CTGGTTAGACCTAGATTTAAACCCCTAAGTAAGATGCTGGGTATAGAACGCTAAGAATT 1320
QY 1256 TTCCCCCAAGGACTCTTGGTTCCTTAAGCCCTTTCGGCTTCGCTTTATGCTCTTCAATAA 1315
Db 1321 TTCCCCCAAGGACTCTTGGTTCCTTAAGCCCTTTCGGCTTCGCTTTATGCTCTTCAATAA 1380
QY 1316 AGTATAAGCCTAACCTTTCGCTAGCTCTAAGGAGAAACCTTTAAACCAAAAGTTTTTAT 1375
Db 1381 AGTATAAGCCTAACCTTTCGCTAGCTCTAAGGAGAAACCTTTAAACCAAAAGTTTTTAT 1440
QY 1376 CATTGAAGACAAATATTGAACACCCCTATTTTGTGGGGATTGAGAAGGGTGAATAGAG 1435
Db 1441 CATTGAAGACAAATATTGAACACCCCTATTTTGTGGGGATTGAGAAGGGTGAATAGAG 1500
QY 1436 CTTGAGACTTTCCTTTGTGGTAGGACTTGGAGAGAAATCCCGTGGACTTTCACTAA 1495
Db 1501 CTTGAGACTTTCCTTTGTGGTAGGACTTGGAGAGAAATCCCGTGGACTTTCACTAA 1560
1496 CCCTCTGACATACCTCCACACCCAGTTGATGGCTTTCCGTAATAAAAGATTGGGATT 1554
1561 CCCTCTGACATACCTCCACACCCAGTTGATGGCTTTCCGTAATAAAAGATTGGGATT 1619
RESULT 3
AX136115 1631-bp DNA linear PAT 30-MAY-2001
LOCUS
DEFINITION Sequence 37 from Patent EP1067182.
ACCESSION AX136115
VERSION AX136115.1 GI:14272523
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 1631)
Ota, T., Isogai, T., Nishikawa, T., Kawai, Y., Sugiyama, T. and
Hayashi, K.
TITLE Secretory protein or membrane protein
JOURNAL Patent: EP 1067182-A 37 10-JAN-2001;
Helix Research Institute (JP)
FEATURES
Location/Qualifiers
1..1631
/organism="Homo sapiens"
/db_xref="taxon:9606"
16..1134
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39700.1"
/db_xref="GI:14272524"
/translation="MAVLAPLIAIVSYVPRLSRWLAQPYLLSALLSRAFLVLRKLPP
LCHGLPTQREDNCPDFDWEVEILMFUSAIVMKNRRSITVEQHIGNIFMFSKVANT
ILFLRLDIRMGLLYTLCLIVLMTCKPPLYMGPEYIKYFNDRKIDIEELERDKRTWIV
EFFANWSDQSFAPYADLSLKYNTGTFNGFVGVGRYDVSRYKYSTPLTKQLP
TLILFOGKEAMRPOIDKKGRAVSFTSEENVIREFNLYORAKLSKAGDNIPE
EOPVAOPQCMQKTRINKILLTAVLPLLSIPGSFHNHKEPAAAFYLCFFPGCDW
GQHAASDFKESRLSGTLQEGLPCCGLFHWSSKREIS"
BASE COUNT 411 a 397 c 378 g 445 t
ORIGIN
Query Match 88.1%; Score 1370.2; DB 6; Length 1631;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 3; Indels 115; Gaps 2;
QY 44 GTTAGGGCGGAAAGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTGGGTGCGG 103
Db 1 GTTAGGGCGGAAAGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTGGGTGCGG 60

Db	1140	AGGGGAAATGTTTCCCTCCCAAGCTTGGGTGAGTGTAACTGCTTATCAGCTATTCA	1199	
Qy	1130	GACATCTCCATGGTTTCTCCATGAACCTCTGTGGTTTCATCATCTCTTCTTATGTTGACCT	1189	
Db	1200	GACATCTCCATGGTTTCTCCATGAACCTCTGTGGTTTCATCATCTCTTCTTATGTTGACCT	1259	
Qy	1190	GCACAGCTTGGTTAGACCTAGATTTAACCCCTAAGCTAAGATGCTGGGTATAGACGCTA	1249	
Db	1260	GCACAGCTTGGTTAGACCTAGATTTAACCCCTAAGCTAAGATGCTGGGTATAGACGCTA	1319	
Qy	1250	AGAAATTTCCCCCAAGGACTCTTGCTTCCCTTAAGCCCTTCTGGCTTCTGTTATGTTCTTC	1309	
Db	1320	AGAAATTTCCCCCAAGGACTCTTGCTTCCCTTAAGCCCTTCTGGCTTCTGTTATGTTCTTC	1379	
Qy	1310	ATTTAAAGTATAAGCTTAACCTTGTCTGCTAGTCTTAAGGAGAAACCTTTAACACAAAGT	1369	
J	1380	ATTTAAAGTATAAGCTTAACCTTGTCTGCTAGTCTTAAGGAGAAACCTTTAACACAAAGT	1439	
Qy	1370	TTTTATCATTTGAAGCAATATTTGAACACCCCTATTTTGTGGGATTCGAGAGGGTGA	1429	
Db	1440	TTTTATCATTTGAAGCAATATTTGAACACCCCTATTTTGTGGGATTCGAGAGGGTGA	1499	
Qy	1430	ATAGAGGCTTGAGACTTTCTCTTGTGTGTAGGACTTGGAGGAGAAATCCCTGGACTTT	1489	
Db	1500	ATAGAGGCTTGAGACTTTCTCTTGTGTGTAGGACTTGGAGGAGAAATCCCTGGACTTT	1559	
Qy	1490	CACTAACCTCTGACATATCTCCACACCCAGTGTGATGCTTTCGTAATAAAGATG	1549	
Db	1560	CACTAACCTCTGACATATCTCCACACCCAGTGTGATGCTTTCGTAATAAAGATG	1619	
Qy	1550	GGATT 1554		
Db	1620	GGATT 1624		
RESULT 4				
AX375964				
LOCUS	AX375964	1620 bp	DNA	linear
DEFINITION	Sequence 31 from Patent WO0168848.			PAT 01-MAR-2002
ACCESSION	AX375964			
VERSION	AX375964.1	GI:19170375		
KEYWORDS	human.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	Baker, K.P., Chen, J., Desnoyers, L., Goddard, A., Godowski, P., J.,			
	Gurney, A.L., Pan, J., Smith, V., Watanabe, C.K., Wood, W.I. and			
	Zhang, Z.			
TITLE	Secreted and transmembrane polypeptides and nucleic acids encoding			
	the same			
JOURNAL	Patent: WO 0168848-A 31.20-SEP-2001;			
Genentech, Inc. (US)				
FEATURES	Location/Qualifiers			
source	1..1620			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
BASE COUNT	407 a 394 c 373 g 442 t			4 others
ORIGIN				
Query Match	88.1%;	Score 1370;	DB 6;	Length 1620;
Best Local Similarity	92.7%;	Pred. No. 0;		
Matches 1494;	Conservative	0;	Mismatches	4;
			Indels	114;
			Gaps	1;
Qy	57	AGATGGGGCTTGGCACCCTTAATGCTCTCGTGTATTCGGTCCGCGACTTTCACGAT	116	
Db	1	AGATGGGGCTTGGCACCCTTAATGCTCTCGTGTATTCGGTCCGCGACTTTCACGAT	60	
Qy	117	GGCTCGGCCAACCTTACTACCTTCTGTGGGCCCTCTCTGCTGCTTCTCTACTCGTGA	176	
Db	61	GGCTCGGCCAACCTTACTACCTTCTGTGGGCCCTCTCTGCTGCTTCTCTACTCGTGA	120	

Qy	177	GGAAATCTCGCGCGCTCTGCCACGGTCTGCCACCGCAACGCGGTAACCCCGTGTG	236	
Db	121	GGAAATCTCGCGCGCTCTGCCACGGTCTGCCACCGCAACGCGGTAACCCCGTGTG	180	
Qy	237	ACTTTGACTGGAGAGAGTGGAGATCCTGATGTTTCTCAGTGCATTTGATGATGAAGA	296	
Db	181	ACTTTGACTGGAGAGAGTGGAGATCCTGATGTTTCTCAGTGCATTTGATGATGAAGA	240	
Qy	297	ACGCGAGATGCCA-----	308	
Db	241	ACGCGAGATGCCA-----	300	
Qy	309	-----	308	
Db	301	CCAACAATCTTTTCCCGCTTGGATATTGCGATGGGCCCTACTTTACATCACACTCT	360	
Qy	309	-----TGTTCTCTGATGAGCTGCAAAACCCCTATATATGGCCCTGAGTATATCAAGT	362	
Db	361	GCATGATGTTCTCTGATGAGCTGCAAAACCCCTATATATGGCCCTGAGTATATCAAGT	420	
Qy	363	ACTTCAATGATAAACCATTTGATGAGAACTAGAACGGGACAAAGAGGGTCACTTGGATTG	422	
Db	421	ACTTCAATGATAAACCATTTGATGAGAACTAGAACGGGACAAAGAGGGTCACTTGGATTG	480	
Qy	423	TGGAGTTCTTTGCCAATTTGCTAATGACTGCCCAATCATTTGCCCTATCTATGCTGACC	482	
Db	481	TGGAGTTCTTTGCCAATTTGCTAATGACTGCCCAATCATTTGCCCTATCTATGCTGACC	540	
Qy	483	TCCTCCCTTAAATACAACTGTACAGGGCTAAATTTTGGAGGGTGGATTTGGAGCGTATA	542	
Db	541	TCCTCCCTTAAATACAACTGTACAGGGCTAAATTTTGGAGGGTGGATTTGGAGCGTATA	600	
Qy	543	CTGATGTTAGTACGCGGTACAAAGTACAGCACATCACCCCTCACCAAGCACTCCCTACCC	602	
Db	601	CTGATGTTAGTACGCGGTACAAAGTACAGCACATCACCCCTCACCAAGCACTCCCTACCC	660	
Qy	603	TGATCTCTTCAAGGTGCAAGGAGGCAATGCGGGGCGCCACAGATTGACAAAGGAC	662	
Db	661	TGATCTCTTCAAGGTGCAAGGAGGCAATGCGGGGCGCCACAGATTGACAAAGGAC	720	
Qy	663	GGCTGTCTCATGGACCTTCTCTGAGGAGAAATGTGATCGAGAAATTTAACTTAAATGAGC	722	
Db	721	GGCTGTCTCATGGACCTTCTCTGAGGAGAAATGTGATCGAGAAATTTAACTTAAATGAGC	780	
Qy	723	TATACAGCGGCGCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCTG	782	
Db	781	TATACAGCGGCGCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCTG	840	
Qy	783	TGGCTTCAACCCCGCCAGTGTGATGGGAAACAAAGAGGATTAATAAGATCCTC	842	
Db	841	TGGCTTCAACCCCGCCAGTGTGATGGGAAACAAAGAGGATTAATAAGATCCTC	900	
Qy	843	ACTTTGGCAGTCTTCTCTCTCTCAATTCAGGCTCTTCCATACCCACAGCAGCTGAG	902	
Db	901	ACTTTGGCAGTCTTCTCTCTCTCAATTCAGGCTCTTCCATACCCACAGCAGCTGAG	960	
Qy	903	GCTGAGCCTTTTATTTATCTTTTCCCTTTGGCTGTGACTGGGTGGGCGGAGCATGAGCT	962	
Db	961	GCTGAGCCTTTTATTTATCTTTTCCCTTTGGCTGTGACTGGGTGGGCGGAGCATGAGCT	1020	
Qy	963	TCTGATTTTAAAGAGCATCTAGGAAATTTGTCAGGCACCTTACAGAAAGCCCTGCCATGC	1022	
Db	1021	TCTGATTTTAAAGAGCATCTAGGAAATTTGTCAGGCACCTTACAGAAAGCCCTGCCATGC	1080	
Qy	1023	TGTCGCCACTCTTTTCACTGGAGCAGAAAGAGATCTCATAGGACGGAGGGGAAATGTT	1082	
Db	1081	TGTCGCCACTCTTTTCACTGGAGCAGAAAGAGATCTCATAGGACGGAGGGGAAATGTT	1140	
Qy	1083	TTCCCTCCAAAGCTTGGGTGAGTGTGTTAACTGCTTATCAGCTATTTCAGACATCTCCATGG	1142	
Db	1141	TTCCCTCCAAAGCTTGGGTGAGTGTGTTAACTGCTTATCAGCTATTTCAGACATCTCCATGG	1200	

QY 964 CTGATTTTAAAGAGGCATCTAGGAAATGTTCAGGCACCTCAGGAAGGCGCTGCATGCT 1023
 |||||
 Db 58670 CTGATTTTAAAGAGGCATCTAGGAAATGTTCAGGCACCTCAGGAAGGCGCTGCATGCT 58611
 |||||
 QY 1024 GTGCCCAACTGTTTCACTGGAGCAAGAAGAGATCTCTATAGGAGGAGGGAATGGTT 1083
 |||||
 Db 58610 GTGCCCAACTGTTTCACTGGAGCAAGAAGAGATCTCTATAGGAGGAGGGAATGGTT 58551
 |||||
 QY 1084 TCCTTCAAGCTTGGGTCAGTGTGTTACTGCTTATCAGCTATTCAGACATCTCCATGCT 1143
 |||||
 Db 58550 TCCTTCAAGCTTGGGTCAGTGTGTTACTGCTTATCAGCTATTCAGACATCTCCATGCT 58491
 |||||
 QY 1144 TTCTCCATGAACCTCTGTGGTTTCATCATCTCTTCTTACCTGACCTGCACAGCTTGGTTA 1203
 |||||
 Db 58490 TTCTCCATGAACCTCTGTGGTTTCATCATCTCTTCTTACCTGACCTGCACAGCTTGGTTA 58431
 |||||
 - 1204 GACCTAGATTAAACCTTAAGTAAGATGCTGGGGTATAGACGCTTAAGAAATTTTCCCCCA 1263
 |||||
 Db 58430 GACCTAGATTAAACCTTAAGTAAGATGCTGGGGTATAGACGCTTAAGAAATTTTCCCCCA 58371
 |||||
 QY 1264 AGACCTCTGCTTCCCTTAAGCCCTCTGGCTTCGTTTATGCTTCATTAAGATATAAG 1323
 |||||
 Db 58370 AGACCTCTGCTTCCCTTAAGCCCTCTGGCTTCGTTTATGCTTCATTAAGATATAAG 58311
 |||||
 QY 1324 CCTAACT 1330
 |||||
 Db 58310 CCTGGCT 58304

RESULT 7
 AC073598
 LOCUS
 DEFINITION Homo sapiens chromosome 10 clone CTC-268N23, WORKING DRAFT
 SEQUENCE, 9 unordered pieces.
 AC073598
 VERSION AC073598.3 GI:20279338
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 112389)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amarantunga,H.C., Are,J.R., Ayalew,M., Banks,F.,
 Barbara,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Dearthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Eathwaite,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
 Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,X., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,

Rives,M., Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savary,G.,
 Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
 Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 112389)
 Worley,K.C.
 Direct Submission
 Submitted (26-JUN-2000) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 112389)
 Worley,K.C.
 Direct Submission
 Submitted (09-MAY-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Apr 24, 2002 this sequence version replaced gi:16118028.

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

Project Information
 Center project name: HMRT
 Center clone name: CTC-268N23

Summary Statistics
 Sequencing vector: M13

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 104470 bases at least Q40

Consensus quality: 107539 bases at least Q30

Consensus quality: 109442 bases at least Q20

Estimated insert size: 111011; sum-of-contigs estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

consists of 9 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

2191: contig of 2191 bp in length

2192: gap of unknown length

2292: contig of 3607 bp in length

5898: gap of unknown length

5999: contig of 5408 bp in length

11407: gap of unknown length

11507: contig of 10233 bp in length

21740: gap of unknown length

21840: contig of 10117 bp in length

31957: gap of unknown length

32057: contig of 14718 bp in length

46775: gap of unknown length

46875: contig of 21344 bp in length

68219: gap of unknown length

68319: contig of 17636 bp in length

85955: gap of unknown length

86055: 112389: contig of 26335 bp in length.

Location/Qualifiers

1..112389

/organism="Homo sapiens"

FEATURES
 source

```
/db_xref="taxon:9606"
/chromosome="10"
/cloname="CTC-268N23"
BASE COUNT 29689 a 25997 c 25893 g 29967 t 843 others
ORIGIN

Query Match 60.1%; Score 934.2; DB 2; Length 112389;
Best Local Similarity 95.6%; Pred. No. 1.9e-281;
Matches 983; Conservative 0; Mismatches 39; Indels 6; Gaps 2;

QY 309 TGTCTCTGATGAGTCAAAACCC-----CCCTATATATGGCCCTGAGTATATCAAGTA 363
DB 19739 TGTCTCTGATGAGTCAAAACCC-----CCCTATATATGGCCCTGAGTATATCAAGTA 19798
QY 364 CTTCAATGATAAAACCAATTGATGAGAACTAGAACGGGACAGAGGGTCACTTGGATTGT 423
DB 19799 CTTCAATGATAAAACCAATTGATGAGAACTAGAACGGGACAGAGGGTCACTTGGATTGT 19858
QY 424 GGAGTTCTTTGCCAATTGCTTAATCACTGCCAATCATTTTGGCCCTATCTATCTGACCT 483
DB 19859 GGAGTTCTTTGCCAATTGCTTAATCACTGCCAATCATTTTGGCCCTATCTATCTGACCT 19918
QY 484 CTCCTCTTAATAACAACGTACAGGGCTAAATTTTGGAGGCTGATGTTGGAGCGCTATAC 543
DB 19919 CTCCTCTTAATAACAACGTACAGGGCTAAATTTTGGAGGCTGATGTTGGAGCGCTATAC 19978
QY 544 TGATGTTAGTACGGGTACAAAGTGTAGACATCACCCCTCCACCAAGCAATCCCTACCT 603
DB 19979 TGATGTTAGTACGGGTACAAAGTGTAGACATCACCCCTCCACCAAGCAATCCCTACCT 20038
QY 604 GATCTGTTCCAGGTGGCAAGGAGCAATCGCGGCCACACATTGACAAAGAGGACG 663
DB 20039 GATCTGTTCCAGGTGGCAAGGAGCAATCGCGGCCACACATTGACAAAGAGGACG 20098
QY 664 GGTCTGCTCATGGACC-TTCTCTGAGGAGAAATGTGATCCGAGAAATTAACCTAAATGAGC 722
DB 20099 GGTCTGCTCATGGACC-TTCTCTGAGGAGAAATGTGATCCGAGAAATTAACCTAAATGAGC 20158
QY 723 TATACAGCGGGCCAAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCGTG 782
DB 20159 TATACAGCGGGCCAAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAGCGTG 20218
QY 783 TGGCTTCACCCCCACACAGTGTACAGTGGGAAACACAGAGGATAAATAGATCCTC 842
DB 20219 TGGCTTCACCCCCACACAGTGTACAGTGGGAAACACAGAGGATAAATAGATCCTC 20278
QY 843 ACTTTGGCAGTGTCTCTCTCTCAATTCAGGCTTTTCCATTAACCAACAGCGCTGAG 902
DB 20279 ACTTTGGCAGTGTCTCTCTCTCAATTCAGGCTTTTCCATTAACCAACAGCGCTGAG 20338
QY 903 GCTGACGCTTTTATTTATTTTCCCTTTTGGCTGTGACTGGTGGGCGAGCATGACGT 962
DB 20339 ACTACAGCCTTTTATTTATTTTCCCTTTTGGCTGTGACTGGTGGGCGAGCATGACGT 20398
QY 963 TCTGATTTTAAAGAGCATCTAGGAAATGTCTAGGACCCCTACAGNAGCGCTGCCATGC 1022
DB 20399 TCTGATTTTAAAGAGCATCTAGGAAATGTCTAGGACCCCTACAGNAGCGCTGCCATGC 20458
QY 1023 TGTGGCAACTGTGTTCACTGGACAGAAAGAGATCTCATAGGAGGAGGGGGAATGGT 1082
DB 20459 TGTGGCAACTGTGTTCACTGGACAGAAAGAGATCTCATAGGAGGAGGGGGAATGGT 20518
QY 1083 TTCCCTCCAGCTTGGGTGAGTGTGTTAACTGCTTATCAGCTATTCAGACATCTCCATGG 1142
DB 20519 TTCCCTCCAGCTTGGGTGAGTGTGTTAACTGCTTATCAGCTATTCAGACATCTCCATGG 20578
QY 1143 TTCTCCATGAACCTCTGGTTTTCATCATCTCTTCTAGTACCTGACACAGCTTGGTT 1202
DB 20579 TTCTCCATGAACCTCTGGTTTTCATCATCTCTTCTAGTACCTGACACAGCTTGGTT 20638
QY 1203 AGACCTTAGATTTAAACCTTAAGTAAAGTATGCTGGGTATAGAACGCTAAAGAAATTTCCCCC 1262
DB 20639 AGACCTTAGATTTAAACCTTAAGTAAAGTATGCTGGGTATAGAACGCTAAAGAAATTTCCCCC 20698
```

```
QY 1263 AAGGACTCTTGTCTCTTAAGCCCTTCTGCTCGTTTATGGTCTTCAATTAAGATATAA 1322
DB 20699 AAGGACTCTTGTCTCTTAAGCCCTTCTGCTCGTTTATGGTCTTCAATTAAGATATAA 20758
QY 1323 GCCTAATCT 1330
DB 20759 GCCTGAGCT 20766

RESULT 8
AP001931
LOCUS Homo sapiens genomic DNA, chromosome 11q clone:RP11-691N7, complete
DEFINITION sequences.
ACCESSION AP001931
VERSION AP001931.5 GI:21218128
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:RP11-691N7.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2000)
REFERENCE 2 (bases 1 to 110116)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (28-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsr.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On May 27, 2002 this sequence version replaced gi:20334325.
FEATURES
location/Qualifiers
1..110116
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-691N7"
BASE COUNT 29082 a 23179 c 24671 g 33184 t

Query Match 55.8%; Score 868; DB 9; Length 110116;
Best Local Similarity 100.0%; Pred. No. 1.2e-260; Indels 0; Gaps 0;
Matches 868; Conservative 0; Mismatches 0;

QY 687 AGGAGATGTGATCCGAGAAATTTAACTTAATGAGCTATACCAGCGGCCCAAGAACTAT 746
DB 20154 AGGAGATGTGATCCGAGAAATTTAACTTAATGAGCTATACCAGCGGCCCAAGAACTAT 20213
QY 747 CAAGAGCTGGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACACAGTGT 806
DB 20214 CAAGAGCTGGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACACAGTGT 20273
QY 807 CAGATGGGAAAACAAGAGATAAATAGATCTCACTTTGGCAGTCTCTCTCTCTG 866
DB 20274 CAGATGGGAAAACAAGAGATAAATAGATCTCACTTTGGCAGTCTCTCTCTCTG 20333
QY 867 TCAATTCAGGCTCTTTCCATAACCAAGCCTGAGCTGCAGCTTTTATTTATGTTTT 926
DB 20334 TCAATTCAGGCTCTTTCCATAACCAAGCCTGAGCTGCAGCTTTTATTTATGTTTT 20393
QY 927 CCCTTTGGCTGTGACTGGGTGGGCGAGCATGCACTTCTGATTTTAAAGAGCATCTAGG 986
DB 20394 CCCTTTGGCTGTGACTGGGTGGGCGAGCATGCACTTCTGATTTTAAAGAGCATCTAGG 20453
QY 987 GAATTTGTCAGGACCCCTACAGGAGGCTGCCATGCTGTGGCAACTGTTTCACTGAGC 1046
```



```

Db      20454 GAATGTCAGGACCCCTACAGGAAGCCCTGCCATGCTGTGGCCAACTGTTTCTACTGGAGC 20513
QY      1047 AAGAAGAGATCTCATAGACGAGGAGGGAATGTTTCCCTCCAAAGCTTGGGTCAAGTGT 1106
Db      20514 AAGAAGAGATCTCATAGACGAGGAGGGAATGTTTCCCTCCAAAGCTTGGGTCAAGTGT 20573
QY      1107 GTTAACTGCTTATCAGTATTCAGACATCTCCATGTTTCTCCATGAAACTCTGTGGTTT 1166
Db      20574 GTTAACTGCTTATCAGTATTCAGACATCTCCATGTTTCTCCATGAAACTCTGTGGTTT 20633
QY      1167 CATCATCTCTTCTAGTACCTGACACGCTTGGTTAGACCTAGATTAAACCTAAGGTA 1226
Db      20634 CATCATCTCTTCTAGTACCTGACACGCTTGGTTAGACCTAGATTAAACCTAAGGTA 20693
QY      1227 AGATCTGGGTATAGAACCCCTAAGAAATTTCCCCCAAGGACTTCTGCTTCCCTTAAGCCC 1286
Db      20694 AGATCTGGGTATAGAACCCCTAAGAAATTTCCCCCAAGGACTTCTGCTTCCCTTAAGCCC 20753
QY      1287 TCTCGCTTCGTTTATGCTCTTCAATAAAGTATAGCCCTAATTTGCTGCTAGTCTCTAA 1346
Db      20754 TCTCGCTTCGTTTATGCTCTTCAATAAAGTATAGCCCTAATTTGCTGCTAGTCTCTAA 20813
QY      1347 GGAGAAACCTTTAACCACAAAGTTTATCATTTGAAGCAATATGAACAACCCCTATT 1406
Db      20814 GGAGAAACCTTTAACCACAAAGTTTATCATTTGAAGCAATATGAACAACCCCTATT 20873
QY      1407 TTGTGGGATTCAGAGGGGTGAATAGAGCTTGAAGCTTCTGCTGCTAGTCTCTAA 1466
Db      20874 TTGTGGGATTCAGAGGGGTGAATAGAGCTTGAAGCTTCTGCTGCTAGTCTCTAA 20933
QY      1467 GGAGAGAAATCCCTGGACTTTCACCTAACCCCTCTGACATCTCCACACCCAGTTGAT 1526
Db      20934 GGAGAGAAATCCCTGGACTTTCACCTAACCCCTCTGACATCTCCACACCCAGTTGAT 20993
QY      1527 GCCTTTCGTAATAAAGATTGGGATT 1554
Db      20994 GCCTTTCGTAATAAAGATTGGGATT 21021

RESULT 9
LOCUS   AP001082
DEFINITION Homo sapiens chromosome 11 clone CMB9-77p23 map 11q12, WORKING
DRAFT SEQUENCE, 32 unordered pieces.
'CESSION
ASTON   AP001082.3 GI:8117229
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE  Homo sapiens DNA, clone:CMB9-77p23.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 130642)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 130,642 genomic DNA of 11q12
Published Only in Database (2000)
2 (bases 1 to 130642)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (25-JAN-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@psc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
On May 30, 2000 this sequence version replaced gi:6997796.
----- Genomic Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@psc.riken.go.jp
----- Project Information
-----

```

```

Center project name: HumDraft11
Center clone name: CMB9-77p23
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 109722 bases at least Q40
Consensus quality: 118358 bases at least Q30
Consensus quality: 123353 bases at least Q20
Insert size: 127542; sum-of-ctrls
Quality coverage: 4.10x in Q20 bases; sum-of-ctrls
-----

```

NOTE: This is a 'working draft' sequence. It currently consists of 32 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1 14868 contig of 14868 bp in length
14969 24185 contig of 9217 bp in length
24286 32497 contig of 8212 bp in length
32598 39406 contig of 6809 bp in length
39507 44898 contig of 5392 bp in length
44999 50122 contig of 5124 bp in length
50223 57989 contig of 7767 bp in length
58090 64088 contig of 5999 bp in length
68189 68556 contig of 4368 bp in length
72259 72259 contig of 3603 bp in length
72360 76692 contig of 4333 bp in length
76793 80690 contig of 3898 bp in length
80791 84253 contig of 3463 bp in length
84354 88683 contig of 4330 bp in length
91721 91721 contig of 2938 bp in length
95301 95301 contig of 3480 bp in length
99413 99413 contig of 4012 bp in length
103118 103118 contig of 3605 bp in length
106147 106147 contig of 2929 bp in length
109453 109453 contig of 3206 bp in length
111611 111611 contig of 2058 bp in length
113460 113460 contig of 1749 bp in length
115886 115886 contig of 2326 bp in length
118771 118771 contig of 2785 bp in length
120623 120623 contig of 1752 bp in length
122472 122472 contig of 1749 bp in length
123753 123753 contig of 1181 bp in length
124856 124856 contig of 1003 bp in length
126517 126517 contig of 1561 bp in length
128115 128115 contig of 1498 bp in length
129271 129271 contig of 1056 bp in length
130642 130642 contig of 1271 bp in length
Sequence updated (02-FEB-2000)
Sequence updated (26-MAY-2000)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 14868: contig of 14868 bp in length
* 14869 14968: gap of 100 bp
* 14969 24185: contig of 9217 bp in length
* 24186 24285: gap of 100 bp
* 32497 32497: contig of 8212 bp in length
* 32498 32597: gap of 100 bp
* 32598 39406: contig of 6809 bp in length
* 39407 39506: gap of 100 bp
* 39507 44898: contig of 5392 bp in length
* 44899 44998: gap of 100 bp
* 44999 50122: contig of 5124 bp in length
* 50123 50222: gap of 100 bp

```



```

Db      52555 GGAGAAACCTTTAACCAACAAAGTTTTTATCATTTGAAGACAATATTGAACAAACCCCTATT 52614
QY      1407 TTGTGGGATTGAGAGGGGTATAGAGGCTTGAGACTTTCCCTTGTGTGTAGGACTT 1466
        |||||||
Db      52615 TTGTGGGATTGAGAGGGGTATAGAGGCTTGAGACTTTCCCTTGTGTGTAGGACTT 52674
        |||||||
QY      1467 GGAGGAGAAATCCCTCGACTTTTCACATAACCTCTGACATACCTCCCAACACCCAGTTGAT 1526
        |||||||
Db      52675 GGAGGAGAAATCCCTCGACTTTTCACATAACCTCTGACATACCTCCCAACACCCAGTTGAT 52734
        |||||||
QY      1527 GCCTTCCGTAATAAAAGATTGGGATT 1554
        |||||||
Db      52735 GCCTTCCGTAATAAAAGATTGGGATT 52762
        |||||||

RESULT 10
309438
CUS
DEFINITION
ACCESSION AC009438
VERSION AC009438.3 GI:9966254
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 158349)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collumore,A.,
Cooke,P., DeArellano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karacas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,N., McEwan,P., McKernan,K., McLaughlin,J.,
Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (22-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 4, 2000 this sequence version replaced gi:7655447.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1359
Center clone name: 77_M17
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 95% of reads
Chemistry: Dye-terminator Big Dye; 95% of reads
Assembly program: Phrap; version 0.960731.
Consensus quality: 138242 bases at least Q40
Consensus quality: 145728 bases at least Q30
Consensus quality: 149650 bases at least Q20
Insert size: 172000; agarose-fp
Insert size: 155949; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; agarose-fp

```

Quality coverage: 4.4 in Q20 base.
 NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
 3651 3750: contig of 3650 bp in length
 3751 4867: contig of 1117 bp in length
 4868 4967: gap of 100 bp
 4968 6757: contig of 1790 bp in length
 6758 6857: gap of 100 bp
 6858 7991: contig of 1134 bp in length
 7992 8091: gap of 100 bp
 8092 9216: contig of 1125 bp in length
 9217 9316: gap of 100 bp
 9317 10384: contig of 1068 bp in length
 10385 10484: gap of 100 bp
 10485 11835: contig of 1351 bp in length
 11836 11935: gap of 100 bp
 11936 13429: contig of 1494 bp in length
 13430 13529: gap of 100 bp
 13530 15092: contig of 1563 bp in length
 15093 15192: gap of 100 bp
 15193 17380: contig of 2188 bp in length
 17381 17480: gap of 100 bp
 17481 20586: contig of 3106 bp in length
 20587 20686: gap of 100 bp
 20687 24958: contig of 4272 bp in length
 24959 25058: gap of 100 bp
 25059 30271: contig of 5213 bp in length
 30272 30371: gap of 100 bp
 30372 35676: contig of 5305 bp in length
 35677 35776: gap of 100 bp
 35777 40640: contig of 4864 bp in length
 40641 40740: gap of 100 bp
 40741 46368: contig of 5628 bp in length
 46369 46468: gap of 100 bp
 46469 53373: contig of 6905 bp in length
 53374 53473: gap of 100 bp
 53474 60571: contig of 7098 bp in length
 60572 60671: gap of 100 bp
 60672 67052: contig of 6381 bp in length
 67053 67152: gap of 100 bp
 67153 75831: contig of 8679 bp in length
 75832 75931: gap of 100 bp
 75932 84118: contig of 8187 bp in length
 84119 84218: gap of 100 bp
 84219 95075: contig of 10857 bp in length
 95076 95175: gap of 100 bp
 95176 110669: contig of 15494 bp in length
 110670 110769: gap of 100 bp
 110770 139814: contig of 29045 bp in length
 139815 139914: gap of 100 bp
 139915 158349: contig of 18435 bp in length.

FEATURES
 source
 1. 158349
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-77M17"
 /clone.lib="RPC1-11 Human Male BAC"
 1. 3650
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 3751. 4867
 /note="assembly_fragment"
 4968. 6757
 misc_feature
 misc_feature
 misc_feature

```
misc_feature      /note="assembly_fragment"  
6858. 7991  
misc_feature      /note="assembly_fragment"  
8092. 9216  
misc_feature      /note="assembly_fragment"  
9317. 10384  
misc_feature      /note="assembly_fragment"  
10485. 11835  
misc_feature      /note="assembly_fragment"  
11936. 13429  
misc_feature      /note="assembly_fragment"  
13530. 15092  
misc_feature      /note="assembly_fragment"  
15193. 17390  
misc_feature      /note="assembly_fragment"  
17481. 20586  
misc_feature      /note="assembly_fragment"  
20687. 24958  
misc_feature      /note="assembly_fragment"  
25059. 30271  
misc_feature      /note="assembly_fragment"  
30372. 35676  
misc_feature      /note="assembly_fragment"  
35777. 40640  
misc_feature      /note="assembly_fragment"  
40741. 46368  
misc_feature      /note="assembly_fragment"  
46469. 53373  
misc_feature      /note="assembly_fragment"  
53474. 60571  
misc_feature      /note="assembly_fragment"  
60672. 67052  
misc_feature      /note="assembly_fragment"  
67153. 75831  
misc_feature      /note="assembly_fragment"  
75932. 84118  
misc_feature      /note="assembly_fragment"  
84219. 95075  
misc_feature      /note="assembly_fragment"  
95176. 110669  
misc_feature      /note="assembly_fragment"  
110770. 139814  
misc_feature      /note="assembly_fragment"  
139915. 158349  
/note="assembly_fragment  
clone_end:T7  
vector_side:right"  
BASE COUNT      40315 a 36712 c 36032 g 42881 t 2409 others  
ORIGIN  
Query Match      55.8%; Score 868; DB 2; Length 158349;  
Best Local Similarity 100.0%; Pred. No. 1.3e-260;  
atches 868; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 687 AGGAGAAATGATCCGAGAAATTAACCTTAAGTACGATACACAGCGGGCCAGAAACTAT 746  
DB 144815 AGGAGAAATGATCCGAGAAATTAACCTTAAGTACGATACACAGCGGGCCAGAAACTAT 144874  
QY 747 CAAGAGCTGGAGACAAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCGCCACACAGTGT 806  
DB 144875 CAAGAGCTGGAGACAAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCGCCACACAGTGT 144934  
QY 807 CAGATGGGGAAACAAAGAGGATAAATAGATCCTCAGCTTTGGCAGTGTCTCTCCTG 866  
DB 144935 CAGATGGGGAAACAAAGAGGATAAATAGATCCTCAGCTTTGGCAGTGTCTCTCCTG 144994  
QY 867 TCAATTCAGGCTCTTCCATACACAGGCTGAGGCTGCAGCCCTTTATTTATGTTTT 926  
DB 144995 TCAATTCAGGCTCTTCCATACACAGGCTGAGGCTGCAGCCCTTTATTTATGTTTT 145054  
QY 927 CCCTTTGGCTGTGACTGTGGGTGGGGCAGCATGAGCTTCTGATTTTAAAGAGGCATCTAGG 986  
DB 145055 CCCTTTGGCTGTGACTGTGGGTGGGGCAGCATGAGCTTCTGATTTTAAAGAGGCATCTAGG 145114
```

```
QY 987 GAATTCCTCAGGACCCCTACAGGAAGCCCTGCCATGCTGTGGCCAACTGTTTCACTGAGC 1046  
DB 145115 GAATTCCTCAGGACCCCTACAGGAAGCCCTGCCATGCTGTGGCCAACTGTTTCACTGAGC 145174  
QY 1047 AAGAAAGAGATCTCATAGGACGAGGGGAAATGGTTTCCTCCAAAGCTTGGGTCACTGT 1106  
DB 145175 AAGAAAGAGATCTCATAGGACGAGGGGAAATGGTTTCCTCCAAAGCTTGGGTCACTGT 145234  
QY 1107 GTTAAGTCTTATCAGCTATTCAGACATCTCCATGGTTTCTCCATGAATCTGTGTTTT 1166  
DB 145235 GTTAAGTCTTATCAGCTATTCAGACATCTCCATGGTTTCTCCATGAATCTGTGTTTT 145294  
QY 1167 CATCATTCCTTCTAGCTGACAGCTGCTGTTAGACCTAGATTTAACCCCTAAGSTA 1226  
DB 145295 CATCATTCCTTCTAGCTGACAGCTGCTGTTAGACCTAGATTTAACCCCTAAGSTA 145354  
QY 1227 AGATGCTGGGTATAGAACGCTAAGAAATTTTCCCCCAAGGACTCTTCTTCTTAAAGCCC 1286  
DB 145355 AGATGCTGGGTATAGAACGCTAAGAAATTTTCCCCCAAGGACTCTTCTTCTTAAAGCCC 145414  
QY 1287 TTCTGGCTTCGTTTATGGTCTTCATTAAGATATAAGCCTAACTTTGTCGCTAGTCTTAA 1346  
DB 145415 TTCTGGCTTCGTTTATGGTCTTCATTAAGATATAAGCCTAACTTTGTCGCTAGTCTTAA 145474  
QY 1347 GGAGAAACCTTTAACCACAAAGTTTTTATCATTTGAAGACAAATATTGAACAAACCCCTATT 1406  
DB 145475 GGAGAAACCTTTAACCACAAAGTTTTTATCATTTGAAGACAAATATTGAACAAACCCCTATT 145534  
QY 1407 TTCTGGGGATTGAGAAGGGGTGAATAGAGGCTTTGAGACTTTTCTTCTGTGTAGGACTTT 1466  
DB 145535 TTCTGGGGATTGAGAAGGGGTGAATAGAGGCTTTGAGACTTTTCTTCTGTGTAGGACTTT 145594  
QY 1467 GGAGGAGAAATCCCTGGACTTTTCACTAACCCCTCTGACATACCTCCACACACCCAGTTGAT 1526  
DB 145595 GGAGGAGAAATCCCTGGACTTTTCACTAACCCCTCTGACATACCTCCACACACCCAGTTGAT 145654  
QY 1527 GCCTTTCCGTAAATAAAGATTTGGGATT 1554  
DB 145655 GCCTTTCCGTAAATAAAGATTTGGGATT 145682  
RESULT 11  
AC090218/c  
LOCUS  
DEFINITION  
Homo sapiens chromosome 11 clone RP11-691N7 map 11, WORKING DRAFT  
SEQUENCE, 17 unordered pieces.  
AC090218.3 GI:14718351  
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.  
VERSION  
KEYWORDS  
SOURCE  
Homo sapiens  
Homo sapiens  
ORGANISM  
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 161397)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barnen, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A.,  
Cammarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collumore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J., S.,  
Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J.,  
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Karakas, A., LaRoque, K., Lamazares, R., Landers, T.,  
Lehoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,  
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,  
McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
```

Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Sougniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Trivisani, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A., and Zody, M.

Direct Submission
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 330 Charles Street, Cambridge, MA 02141, USA
On July 14, 2001 this sequence version replaced gi:13940674.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L12248

Center clone name: 691_N_7

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 155030 bases at least Q40

Consensus quality: 157856 bases at least Q30

Consensus quality: 159050 bases at least Q20

Insert size: 157000; agarose-fp

Insert size: 159797; sum-of-contigs

Quality coverage: 12.0 in Q20 bases; agarose-fp

Quality coverage: 11.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 7057: contig of 7057 bp in length

* 7058 7157: gap of 100 bp

* 7158 7519: contig of 362 bp in length

* 7520 7619: gap of 100 bp

* 7620 8839: contig of 1220 bp in length

* 8840 8939: gap of 100 bp

* 8940 9986: contig of 1047 bp in length

* 9987 10086: gap of 100 bp

* 10087 11612: contig of 1526 bp in length

* 11613 11712: gap of 100 bp

* 11713 13287: contig of 1575 bp in length

* 13288 13387: gap of 100 bp

* 13388 14435: contig of 1048 bp in length

* 14436 14535: gap of 100 bp

* 14536 16619: contig of 2084 bp in length

* 16620 16719: gap of 100 bp

* 16720 17857: contig of 1138 bp in length

* 17858 17957: gap of 100 bp

* 17958 19972: contig of 2015 bp in length

* 19973 20072: gap of 100 bp

* 20073 22132: contig of 2060 bp in length

* 22133 22232: gap of 100 bp

* 22233 58999: contig of 36667 bp in length

* 58900 58999: gap of 100 bp

* 59000 70617: contig of 11618 bp in length

* 70618 70717: gap of 100 bp

* 70718 89521: contig of 18804 bp in length

* 89522 89621: gap of 100 bp

* 89622 118253: contig of 28632 bp in length

* 118254 118353: gap of 100 bp

* 118354 155823: contig of 37470 bp in length

* 155824 155924: gap of 100 bp

* 155924 161397: contig of 5474 bp in length.

* 161397 161397: contig of 5474 bp in length.

* 161397 161397: contig of 5474 bp in length.

* 161397 161397: contig of 5474 bp in length.

* 161397 161397: contig of 5474 bp in length.

* 161397 161397: contig of 5474 bp in length.

* 161397 161397: contig of 5474 bp in length.

* 155824 155923: gap of 100 bp

* 155924 161397: contig of 5474 bp in length.

FEATURES

source

Location/Qualifiers

1..161397

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone_lib="RPC1-11 Human Male BAC"

1..7057

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

7158..7519

/note="assembly_fragment"

7620..8839

/note="assembly_fragment"

8940..9986

/note="assembly_fragment"

10087..11612

/note="assembly_fragment"

11713..13287

/note="assembly_fragment"

13388..14435

/note="assembly_fragment"

14536..16619

/note="assembly_fragment"

16720..17857

/note="assembly_fragment"

17958..19972

/note="assembly_fragment"

20073..22132

/note="assembly_fragment"

22233..58999

/note="assembly_fragment"

59000..70617

/note="assembly_fragment"

70718..89521

/note="assembly_fragment"

89622..118253

/note="assembly_fragment"

118354..155823

/note="assembly_fragment"

155924..161397

/note="assembly_fragment"

clone_end:T7

vector_side:right"

BASE COUNT 43394 a 34530 c 35670 g 46202 t 1601 others

ORIGIN

Query Match 55.8%; Score 868; DB 2; Length 161397;

Best Local Similarity 100.0%; Pred. No. 1.3e-260;

Matches 868; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 687 AGGAGATGTGATCCGAGAAATTAACCTTAATAGCTATACAGCGGCCCAAGAACTAT 746

Db 36543 AGGAGATGTGATCCGAGAAATTAACCTTAATAGCTATACAGCGGCCCAAGAACTAT 36484

Qy 747 CAAAGGCTGGAGACAATATCCCTGAGGAGAGCCTGTGGCTTCAACCCCAACACAGTGT 806

Db 36483 CAAAGGCTGGAGACAATATCCCTGAGGAGAGCCTGTGGCTTCAACCCCAACACAGTGT 36424

Qy 807 CAGATGGGAAACAAGAAGTAATAAGATCCTCACTTTGGGAGTCTTCTCTCTCTG 866

Db 36423 CAGATGGGAAACAAGAAGTAATAAGATCCTCACTTTGGGAGTCTTCTCTCTCTG 36364

Qy 867 TCATTCAGGCTCTTTCCATAACCAAGCCTGAGGCTGAGGCTTTATTTATGTTTT 926

Db 36363 TCATTCAGGCTCTTTCCATAACCAAGCCTGAGGCTGAGGCTTTATTTATGTTTT 36304

Qy 927 CCCTTGGCTGTACTGGTGGGCGACATGCAGCTTCTGATTTAAAGAGGCATCTAGG 986

Db 927 CCCTTGGCTGTACTGGTGGGCGACATGCAGCTTCTGATTTAAAGAGGCATCTAGG 986

```

Db      36303 CCCTTTGGCTGTGACTGGCTGGGCGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGG 36244
QY      987  GAATTCCTAGGCACCTACAGGAGCCCTGCCATGCTGTGGCCAACTGTTTCACTGAGC 1046
Db      36243 GAATTCCTAGGCACCTACAGGAGCCCTGCCATGCTGTGGCCAACTGTTTCACTGAGC 36184
QY      1047 AAGAAGAGATCTCATAGGACGGAGGGGAAATGTTTCCCTCCAAAGCTTGGCTCAGTGT 1106
Db      36183 AAGAAGAGATCTCATAGGACGGAGGGGAAATGTTTCCCTCCAAAGCTTGGCTCAGTGT 36124
QY      1107 GTTAAGCTTATCAGCTATTCAGACATCTCCATGTTTCTCATGAACTCTGTGGTTT 1166
Db      36123 GTTAAGCTTATCAGCTATTCAGACATCTCCATGTTTCTCATGAACTCTGTGGTTT 36064
QY      1167 CATCATCTCTTCTTAGTTCACCTGCACAGCTTGGTTAGACCTAGATTTAAACCTTAAGGTA 1226
Db      36063 CATCATCTCTTCTTAGTTCACCTGCACAGCTTGGTTAGACCTAGATTTAAACCTTAAGGTA 36004
QY      1227 AGATGCTGGGGTATAGAAGCTTAAGAAATTTCCCCCAAGGACTTTCCTTCCCTTAAGCCC 1286
Db      36003 AGATGCTGGGGTATAGAAGCTTAAGAAATTTCCCCCAAGGACTTTCCTTCCCTTAAGCCC 35944
QY      1287 TTCTGGCTTCGTTTATGGTCTTCATTAAAGTATAAGCCTAACTTGTTCGCTAGTCTCTAA 1346
Db      35943 TTCTGGCTTCGTTTATGGTCTTCATTAAAGTATAAGCCTAACTTGTTCGCTAGTCTCTAA 35884
QY      1347 GGAGAAACCTTTAACCACAAAGTTTTTATCATTTGAAGACATATTGAACACCCCTATT 1406
Db      35883 GGAGAAACCTTTAACCACAAAGTTTTTATCATTTGAAGACATATTGAACACCCCTATT 35824
QY      1407 TTGTGGGGATTGAAGAGGGGTGAATAGAGGCTTGAGACTTTCCTTGTGTGGTAGGACTT 1466
Db      35823 TTGTGGGGATTGAAGAGGGGTGAATAGAGGCTTGAGACTTTCCTTGTGTGGTAGGACTT 35764
QY      1467 GGAGGAGAAATCCCTGGACTTTCACATAACCCCTGACATACCTCCACACCCAGCTGTAT 1526
Db      35763 GGAGGAGAAATCCCTGGACTTTCACATAACCCCTGACATACCTCCACACCCAGCTGTAT 35704
QY      1527 GCCTTTCCTTAATAAAGATTGGGATT 1554
Db      35703 GCCTTTCCTTAATAAAGATTGGGATT 35676

RESULT 12
AC092707
LOCUS   AC092707
DEFINITION Homo sapiens chromosome 11 clone RP11-679G21 map 11, WORKING DRAFT
SEQUENCE, 9 unordered pieces.
ACCESSION AC092707
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE  Homo sapiens.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 170540)
    Birren.B., Linton.L., Nusbaum.C. and Lander.B.
  Unpublished
  2 (bases 1 to 170540)
    Birren.B., Linton.L., Nusbaum.C., Lander.E., Allen.A., Allen.N.,
    Anderson.S., Barna.N., Bastien.V., Boguslavskiy.L., Boukhgalter.B.,
    Brown.A., Camarata.J., Campopiano.A., Chang.J., Chazaro.B.,
    Choquel.Y., Colangelo.M., Collins.S., Collamore.A., Cook.A.,
    Cooke.P., DeArrellano.K., Dewar.K., Diaz.J.S., Dodge.S., Faro.S.,
    Ferreira.P., Fitzhugh.W., Gage.D., Galagan.J., Gardyna.S.,
    Ginde.S., Gord.S., Goyette.M., Graham.L., Grand-Pierre.N.,
    Hagos.B., Heaford.A., Horton.L., Hulme.W., Iliev.I., Johnson.R.,
    Jones.C., Kamat.A., Kartas.A., Kells.C., Larocque.K.,
    Lamazares.R., Lander.B., Lander.T., Lehoczyk.J., Levine.R., Liu.G.,
    MacLean.C., Macdonald.P., Major.J., Marquis.N., Matthews.C.,
    McCarthy.M., McEwan.P., McKernan.K., McPheeters.R., Meldrum.J.,
    Meneus.L., Mihova.T., Mlenga.V., Murphy.T., Naylor.J., Nguyen.C.,

```

```

Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta.R., Rieback,M., Riley,R., Rise.C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
Seaman,S., Severy,P., Sougniez,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody.M.
Direct Submission
Submitted (20-JUL-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12706
Center clone name: 679_G_21
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 166430 bases at least Q40
Consensus quality: 166426 bases at least Q30
Consensus quality: 169078 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 169740; sum-of-contigs
Quality coverage: 9.3 in Q20 bases; agarose-fp
Quality coverage: 8.3 in Q20 bases; sum-of-contigs
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

* 1
* 744 843: gap of 100 bp
* 844 1842: contig of 999 bp in length
* 1843 1942: gap of 100 bp
* 1943 4117: contig of 2175 bp in length
* 4118 4217: gap of 100 bp
* 4218 9901: contig of 5684 bp in length
* 9902 10001: gap of 100 bp
* 10002 22030: contig of 12029 bp in length
* 22031 22130: gap of 100 bp
* 22131 77732: contig of 55602 bp in length
* 77733 77832: gap of 100 bp
* 77833 110109: contig of 32277 bp in length
* 110110 110209: gap of 100 bp
* 110210 150387: contig of 40178 bp in length
* 150388 150487: gap of 100 bp
* 150488 170540: contig of 20053 bp in length.

```

```

FEATURES
source
  1..170540
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11"
    /clone="RP11-679G21"
    /clone_lib="RPC1-11 Human Male BAC"
  1..743
    /note="assembly fragment
    clone_end:SP6
    vector_side:left"
misc_feature
  844..1842

```

```
/note="assembly_fragment"
1943. 4117
/note="assembly_fragment"
4218. 9901
/note="assembly_fragment"
10002. 22030
/note="assembly_fragment"
22131. 77732
/note="assembly_fragment"
77833. 110109
/note="assembly_fragment"
110210. 150387
/note="assembly_fragment"
150488. 170540
/note="assembly_fragment
vector_side:right"
vector_end:T7
vector_size:right"
45658 t 806 others
ORIGIN
Query Match 55.8%; Score 868; DB 2; Length 170540;
Best Local Similarity 100.0%; Pred. No. 1.3e-260;
Matches 868; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 687 AGGAGATGTGATCCGAGATTTAACTAAATGAGCTATACAGCGGCCCAAGAACTAT 746
Db 63433 AGGAGATGTGATCCGAGATTTAACTAAATGAGCTATACAGCGGCCCAAGAACTAT 63492

QY 747 CAAGAGCTGGAGACAATATCCCTGAGGAGCGCTGTGGCTTCAACGCCACCACAGTGT 806
Db 63493 CAAGAGCTGGAGACAATATCCCTGAGGAGCGCTGTGGCTTCAACGCCACCACAGTGT 63552

QY 807 CAGATGGGGAACAAGAAGGATAAATAAGATCCTCACTTTGGCAGTGCTTCTCTCTCTG 866
Db 63553 CAGATGGGGAACAAGAAGGATAAATAAGATCCTCACTTTGGCAGTGCTTCTCTCTCTG 63612

QY 867 TCAATTCAGCGCTTTTCCATACCAAGCGCTGAGCTGAGCTTATTTATTTATTTT 926
Db 63613 TCAATTCAGCGCTTTTCCATACCAAGCGCTGAGCTGAGCTTATTTATTTATTTT 63672

QY 927 CCCTTTGGCTGTGACTGGGTGGGCGAGCATGCAGCTTCTGTATTTAAAGAGCGCATCTAGG 986
Db 63673 CCCTTTGGCTGTGACTGGGTGGGCGAGCATGCAGCTTCTGTATTTAAAGAGCGCATCTAGG 63732

QY 987 GAATTTGTCAGCCACCTACAGGAAGGCTGCCATGCTGTGGCCAACTGTTTCACTGGAGC 1046
Db 63733 GAATTTGTCAGCCACCTACAGGAAGGCTGCCATGCTGTGGCCAACTGTTTCACTGGAGC 63792

QY 1047 AAGAAAGAGATCTCATAGGACGGGGGAATGTTTCCCTCCAAGCTTGGGTCAGNGT 1106
Db 63793 AAGAAAGAGATCTCATAGGACGGGGGAATGTTTCCCTCCAAGCTTGGGTCAGNGT 63852

QY 1107 GTTAACCTGTTATCAGCTATTACAGATCTCCATGGTTTCTCCATGAACCTCTGTGGTTT 1166
Db 63853 GTTAACCTGTTATCAGCTATTACAGATCTCCATGGTTTCTCCATGAACCTCTGTGGTTT 63912

QY 1167 CATCATCTCTTCTAGTTGACCTGCACAGCTGTGTTAGACCTAGATTTAAACCCCTAAGGTA 1226
Db 63913 CATCATCTCTTCTAGTTGACCTGCACAGCTGTGTTAGACCTAGATTTAAACCCCTAAGGTA 63972

QY 1227 AGATGCTGGGTATAGAACGCTAGAAATTTTCCCCCAAGGACTCTGCTTCCCTTAAGCCC 1286
Db 63973 AGATGCTGGGTATAGAACGCTAGAAATTTTCCCCCAAGGACTCTGCTTCCCTTAAGCCC 64032

QY 1287 TTCTGGCTTCTGTTATGCTTCTTAAAGCTATAAGCTAACTTTGTCGCTAGTCTCTAA 1346
Db 64033 TTCTGGCTTCTGTTATGCTTCTTAAAGCTATAAGCTAACTTTGTCGCTAGTCTCTAA 64092

QY 1347 GGAGAAACCTTTAACCAACAAGTTTATCATTTGAAGACATATTTGAACACCCCTATT 1406
Db 64093 GGAGAAACCTTTAACCAACAAGTTTATCATTTGAAGACATATTTGAACACCCCTATT 64152

QY 1407 TTGTGGGATTTGAGAGGGGTGAATAGAGGCTTTCCTTTCCTTGTGTGGTAGGACTT 1466

Db 64153 TTGTGGGATTTGAGAGGGGTGAATAGAGGCTTTCCTTTCCTTGTGTGGTAGGACTT 64212
QY 1467 GGAGAGAAATCCCTCGACTTTTCACTAACCTCTGACATACCTCCACACCCAGTTGAT 1526
Db 64213 GGAGAGAAATCCCTCGACTTTTCACTAACCTCTGACATACCTCCACACCCAGTTGAT 64272
QY 1527 GGCITTCGTAATAAAAAAGATTGGGATT 1554
Db 64273 GGCITTCGTAATAAAAAAGATTGGGATT 64300

RESULT 13
AC021522/c
LOCUS AC021522
DEFINITION Homo sapiens clone RP11-335H5, WORKING DRAFT SEQUENCE, 14 unordered
pieces.
AC021522
VERSION AC021522.5 GI:14971311
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 199992)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
2 (bases 1 to 199992)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castele,A.,
Choepeil,Y., Collangelo,M., Collins,S., Collamore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferrelina,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,
Macdonald,P., Markus,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 20, 2001 this sequence version replaced gi:14269795.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5504
Center clone name: 335_H_5
----- Summary Statistics
Sequencing vector: M13; M7815; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 192697 bases at least Q40
Consensus quality: 196336 bases at least Q30
Consensus quality: 197632 bases at least Q20
Insert size: 193000; agarose-fp
Insert size: 198692; sum-of-contents
Quality coverage: 8.3 in Q20 bases; agarose-fp
Quality coverage: 8.1 in Q20 bases.
```

* NOTE: This is a 'working draft' sequence. It currently consists of 14 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 17413: contig of 17413 bp in length
* 17414 17513: gap of 100 bp
* 17514 18244: contig of 731 bp in length
* 18245 18344: gap of 100 bp
* 18345 19084: contig of 740 bp in length
* 19085 19184: gap of 100 bp
* 19185 20416: contig of 1232 bp in length
* 20417 20516: gap of 100 bp
* 20517 21625: contig of 1109 bp in length
* 21626 21725: gap of 100 bp
* 21726 102938: contig of 81213 bp in length
* 102939 103038: gap of 100 bp
* 103039 104905: contig of 1867 bp in length
* 104906 105005: gap of 100 bp
* 105006 108216: contig of 3211 bp in length
* 108217 108316: gap of 100 bp
* 108317 112875: contig of 4559 bp in length
* 112876 112975: gap of 100 bp
* 112976 122676: contig of 9701 bp in length
* 122677 122776: gap of 100 bp
* 122777 138821: contig of 16045 bp in length
* 138822 138921: gap of 100 bp
* 138922 168907: contig of 29986 bp in length
* 168908 169007: gap of 100 bp
* 169008 196861: contig of 27854 bp in length
* 196862 196961: gap of 100 bp
* 196962 199992: contig of 3031 bp in length.

```

FEATURES

source

```

1. 199992
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-335H5"
/clone_lib="RPC1-11 Human Male BAC"

```

misc_feature

```

1. 17413
/notes="assembly_fragment"
clone_end:SP6
vector_side:left

```

misc_feature

```

17514..18244
/notes="assembly_fragment"

```

misc_feature

```

18345..19084
/notes="assembly_fragment"

```

misc_feature

```

19185..20416
/notes="assembly_fragment"

```

misc_feature

```

20517..21625
/notes="assembly_fragment"

```

misc_feature

```

21726..102938
/notes="assembly_fragment"

```

misc_feature

```

103039..104905
/notes="assembly_fragment"

```

misc_feature

```

105006..108216
/notes="assembly_fragment"

```

misc_feature

```

108317..112875
/notes="assembly_fragment"

```

misc_feature

```

112976..122676
/notes="assembly_fragment"

```

misc_feature

```

122777..138821
/notes="assembly_fragment"

```

misc_feature

```

138922..168907
/notes="assembly_fragment"

```

misc_feature

```

169008..196861
/notes="assembly_fragment"

```

misc_feature

```

196862..199992
/notes="assembly_fragment"
clone_end:T7
vector_side:right

```

BASE COUNT

54206 a 46381 c 45637 g 52463 t 1305 others

ORIGIN

Query Match 55.8%; Score 868; DB 2; Length 199992;

Best Local Similarity 100.0%; Pred. No. 1.3e-260;

Matches 868; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 687 AGGAGATGTGATCCGAGAAATTTAACTTAATAGCTATACACGCGGCCCAAGAACTAT 746
|||||
DB 183496 AGGAGATGTGATCCGAGAAATTTAACTTAATAGCTATACACGCGGCCCAAGAACTAT 183437
QY 747 CAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCAACACAGTGT 806
|||||
DB 183436 CAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCAACACAGTGT 183377
QY 807 CAGATGGGAAACAAGAAGGATAAATAAGATCCTACTTTGGCAGTGCTTCTCTCTCTG 866
|||||
DB 183376 CAGATGGGAAACAAGAAGGATAAATAAGATCCTACTTTGGCAGTGCTTCTCTCTCTG 183317
QY 867 TCAATTCAGGCTCTTTCCATAACCAAGCCTCAGGCTGCAGCCTTTATTTATGTTTT 926
|||||
DB 183316 TCAATTCAGGCTCTTTCCATAACCAAGCCTCAGGCTGCAGCCTTTATTTATGTTTT 183257
QY 927 CCCTTTGGCTGTACTGGGTGGGCGCAGCATGCAGCTTCTTGATTTAAAGAGGCATCTAGG 986
|||||
DB 183256 CCCTTTGGCTGTACTGGGTGGGCGCAGCATGCAGCTTCTTGATTTAAAGAGGCATCTAGG 183197
QY 987 GAATTGTCAGGCCCTACAGGAAGGCTGCCATGCTGTGCCCAACTGTTTCACTGGAGC 1046
|||||
DB 183196 GAATTGTCAGGCCCTACAGGAAGGCTGCCATGCTGTGCCCAACTGTTTCACTGGAGC 183137
QY 1047 AAGAAAGAGATCTCATAGGACGGGGAATGGTTTCCCTCCAAGCTTGGGTCAGTGT 1106
|||||
DB 183136 AAGAAAGAGATCTCATAGGACGGGGAATGGTTTCCCTCCAAGCTTGGGTCAGTGT 183077
QY 1107 GTTAACCTCTATCAGCTATTTCAGACATCTCCATGGTTTCTCCATGAACCTCTGTGGTTT 1166
|||||
DB 183076 GTTAACCTCTATCAGCTATTTCAGACATCTCCATGGTTTCTCCATGAACCTCTGTGGTTT 183017
QY 1167 CATCATTCCTTCTTAGTTAGCTGCACAGCTTGGTTAGACCTAGATTTAACCCCTAAGGTA 1226
|||||
DB 183016 CATCATTCCTTCTTAGTTAGCTGCACAGCTTGGTTAGACCTAGATTTAACCCCTAAGGTA 182957
QY 1227 AGATGCTGGGTATAGAACCTTAAGAATTTTCCCAAGGACTCTTCTTCTTAAAGCCC 1286
|||||
DB 182956 AGATGCTGGGTATAGAACCTTAAGAATTTTCCCAAGGACTCTTCTTCTTAAAGCCC 182897
QY 1287 TTCTGGCTTCGTTTATGGTCTTCATTAAGATATAAGCCTAACTTTTGTGCTAGTCTCTAA 1346
|||||
DB 182896 TTCTGGCTTCGTTTATGGTCTTCATTAAGATATAAGCCTAACTTTTGTGCTAGTCTCTAA 182837
QY 1347 GGAGAAACCTTTTAAACCAAGTTTTTATCATTAAGACAAATATTGAACACCCCTATT 1406
|||||
DB 182836 GGAGAAACCTTTTAAACCAAGTTTTTATCATTAAGACAAATATTGAACACCCCTATT 182777
QY 1407 TTGTGGGATTGAGAGGGGTGATAGAGCTTGAGACTTTCCTTTCTGTGGTAGGACTT 1466
|||||
DB 182776 TTGTGGGATTGAGAGGGGTGATAGAGCTTGAGACTTTCCTTTCTGTGGTAGGACTT 182717
QY 1467 GGAGGAGAAATCCCTCGACTTTTCACTAACCTCTGACATCTCCCAACACCCAGTGTAT 1526
|||||
DB 182716 GGAGGAGAAATCCCTCGACTTTTCACTAACCTCTGACATCTCCCAACACCCAGTGTAT 182657
QY 1527 GCCTTTCCGTAATAAAGATTTGGGATT 1554
|||||
DB 182656 GCCTTTCCGTAATAAAGATTTGGGATT 182629
|||||

```

RESULT 14

AP000727/c

LOCUS

Homo sapiens chromosome 11 clone RP11-679G21 map 11q12, WORKING

DEFINITION DRAFT SEQUENCE, 40 unordered pieces.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

AP000727
AP000727.2 GI-8118896
HTG; HTGS-PHASE1; HTGS-DRAFT.
Homo sapiens DNA, clone:RP11-679G21.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 153394)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 153,394 genomic DNA of llql2
Published Only in DataBase (1999)
2 (bases 1 to 153394)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (16-NOV-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp)
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
On May 31, 2000 this sequence version replaced gi:6997582.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP11-679G21
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 133607 bases at least Q40
Consensus quality: 142217 bases at least Q30
Consensus quality: 146720 bases at least Q20
Insert size: 149494; sum-of-contigs
Quality coverage: 4.03x in Q20 bases; sum-of-contigs

COMMENT

NOTE: This is a 'working draft' sequence. It currently consists of 40 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 11871: contig of 11871 bp in length
11972 11971: gap of 100 bp
11972 22347: contig of 10376 bp in length
22348 22447: gap of 100 bp
22448 33540: contig of 11093 bp in length
33541 33640: gap of 100 bp
33641 41577: contig of 7937 bp in length
41578 41677: gap of 100 bp
41678 45975: contig of 4298 bp in length
45976 46075: gap of 100 bp
46076 53127: contig of 7052 bp in length
53128 53227: gap of 100 bp
53228 60314: contig of 7087 bp in length
60315 60414: gap of 100 bp
60415 65544: contig of 5130 bp in length
65545 65644: gap of 100 bp
65645 71027: contig of 5383 bp in length
71028 71127: gap of 100 bp
71128 75422: contig of 4295 bp in length
75423 75522: gap of 100 bp
75523 79274: contig of 3752 bp in length
79275 79374: gap of 100 bp
79375 84398: contig of 5024 bp in length
84399 84498: gap of 100 bp
84499 88990: contig of 4492 bp in length
88991 89090: gap of 100 bp
89091 93436: contig of 4346 bp in length
93437 93536: gap of 100 bp
93537 96655: contig of 3119 bp in length
96656 96755: gap of 100 bp
96756 99175: contig of 2420 bp in length
99176 99275: gap of 100 bp
99276 102659: contig of 3384 bp in length
102660 102759: gap of 100 bp
102760 105824: contig of 3065 bp in length
105825 105924: gap of 100 bp
105925 109361: contig of 3437 bp in length
109362 109461: gap of 100 bp
109462 112303: contig of 2842 bp in length
112304 112403: gap of 100 bp
112404 115484: contig of 3081 bp in length
115485 115584: gap of 100 bp
115585 118972: contig of 3388 bp in length
118973 119072: gap of 100 bp
119073 121620: contig of 2548 bp in length
121621 121720: gap of 100 bp
121721 124322: contig of 2602 bp in length
124323 124422: gap of 100 bp
124423 127088: contig of 2666 bp in length


```

* 127089 127188: gap of 100 bp
* 127189 130306: contig of 3118 bp in length
* 130307 130406: gap of 100 bp
* 130407 131809: contig of 1403 bp in length
* 131810 131909: gap of 100 bp
* 131910 133740: contig of 1831 bp in length
* 133741 133840: gap of 100 bp
* 133841 136109: contig of 2269 bp in length
* 136110 136209: gap of 100 bp
* 136210 137539: contig of 1330 bp in length
* 137540 137639: gap of 100 bp
* 137640 139390: contig of 1751 bp in length
* 139391 139490: gap of 100 bp
* 139491 140939: contig of 1449 bp in length
* 140940 141039: gap of 100 bp
* 141040 142493: contig of 1454 bp in length
* 142494 142593: gap of 100 bp
* 142594 144234: contig of 1641 bp in length
* 144235 144334: gap of 100 bp
* 144335 145850: contig of 1516 bp in length
* 145851 145950: gap of 100 bp
* 145951 147405: contig of 1455 bp in length
* 147406 147505: gap of 100 bp
* 147506 149002: contig of 1497 bp in length
* 149003 149102: gap of 100 bp
* 149103 150523: contig of 1421 bp in length
* 150524 150623: gap of 100 bp
* 150624 152235: contig of 1612 bp in length
* 152236 152335: gap of 100 bp
* 152336 153394: contig of 1059 bp in length.

```

FEATURES

Location/Qualifiers

```

source
  1..153394
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11q12"
    /clone="RP11-679G21"
  misc_feature
    1..11871
      /note="assembly_fragment"
  misc_feature
    11972..22347
      /note="assembly_fragment"
  misc_feature
    22448..33540
      /note="assembly_fragment"
  misc_feature
    33641..41577
      /note="assembly_fragment"
  misc_feature
    41678..45975
      /note="assembly_fragment clone_end:T7 vector_side:left"
  misc_feature
    46076..53127
      /note="assembly_fragment"
  misc_feature
    53228..60314
      /note="assembly_fragment"

```

ery Match 55.0%; Score 856; DB 2: Length 153394;

est Local Similarity 99.9%; Pred. No. 7.4e-257;

Matches 867; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

QY 687 AGGAAATGTGATCCGAGATTTAACTTAATGAGCTATACAGCGGGCCCAAGAACTAT 746
|||||
DB 11478 AGGAAATGTGATCCGAGATTTAACTTAATGAGCTATACAGCGGG-CAAGAACTAT 11420
|||||
QY 747 CAAAGCTCGAGACAATATCCCTGAGGAGCGCTGTGGCTTCAACCCCAACACAGTGT 806
|||||
DB 11419 CAAAGCTCGAGACAATATCCCTGAGGAGCGCTGTGGCTTCAACCCCAACACAGTGT 11360
|||||
QY 807 CAGATGGGAAACAAAGAGGATAAATAGATCCTCAGTGTTCCTCTCCTG 866
|||||
DB 11359 CAGATGGGAAACAAAGAGGATAAATAGATCCTCAGTGTTCCTCTCCTG 11300
|||||
QY 867 TCAATTCAGGCTCTTTCCATACACCAAGCGTGGAGCGCTTTTATTTATGTTTT 926
|||||
DB 11299 TCAATTCAGGCTCTTTCCATACACCAAGCGTGGAGCGCTTTTATTTATGTTTT 11240
|||||
QY 927 CCCTTTGGCTGTGACTGTGGTGGGAGCATGCGAGCTCTGATTTTAAAGAGGCATCTAGG 986

```

```

DB 11239 CCCTTTGGCTGTGACTGTGGTGGGAGCATGCGAGCTCTGATTTTAAAGAGCATCTAGG 11180
QY 987 GAATTTGACGGCACCCTACAGGAAGGCTGCCATGCTGTGGCAACTGTTTCACTGGAGC 1046
|||||
DB 11179 GAATTTGACGGCACCCTACAGGAAGGCTGCCATGCTGTGGCAACTGTTTCACTGGAGC 11120
QY 1047 AAGAAAGAGATCTCATAGGAGCGGGGAAATGTTTCCCTCCAAGCTTGGGTCACTGT 1106
|||||
DB 11119 AAGAAAGAGATCTCATAGGAGCGGGGAAATGTTTCCCTCCAAGCTTGGGTCACTGT 11060
QY 1107 GTTAACTGCTTATCAGCTATTTCAGACATCTCCATGTTTCTCCATGAACCTCTCTGTTTT 1166
|||||
DB 11059 GTTAACTGCTTATCAGCTATTTCAGACATCTCCATGTTTCTCCATGAACCTCTCTGTTTT 11000
QY 1167 CATCATTCCTTCTTAGTTGACCTGTCACAGCTTGGTTAGACCTAGATTTAACCCCTAAGGTA 1226
|||||
DB 10999 CATCATTCCTTCTTAGTTGACCTGTCACAGCTTGGTTAGACCTAGATTTAACCCCTAAGGTA 10940
QY 1227 AGATGCTGGGTATAGAACGCTTAAGAAATTTTCCCCCAAGGACTCTTCTCTTCCCTTAAGCCC 1286
|||||
DB 10939 AGATGCTGGGTATAGAACGCTTAAGAAATTTTCCCCCAAGGACTCTTCTCTTCCCTTAAGCCC 10880
QY 1287 TTCTGGCTTCGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 1346
|||||
DB 10879 TTCTGGCTTCGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 10820
QY 1347 GGAGAAACCTTTAACCAACAAAGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 1406
|||||
DB 10819 GGAGAAACCTTTAACCAACAAAGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 10760
QY 1407 TTCTGGGATTGAGAGGGGTGATAGAGGCTTGTAGAGCTTTCCTTTCTGTTGGTAGGACTT 1466
|||||
DB 10759 TTCTGGGATTGAGAGGGGTGATAGAGGCTTGTAGAGCTTTCCTTTCTGTTGGTAGGACTT 10700
QY 1467 GGAGGAGAAATCCCTGGACTTTTCACTAACCTCTGACATACCTCCCAACACCCAGTTTGTAT 1526
|||||
DB 10699 GGAGGAGAAATCCCTGGACTTTTCACTAACCTCTGACATACCTCCCAACACCCAGTTTGTAT 10640
QY 1527 GGCTTTCCGTAATAAAAGATTGGGATT 1554
|||||
DB 10639 GGCTTTCCGTAATAAAAGATTGGGATT 10612

```

RESULT 15

AC013381

LOCUS

DEFINITION

SEQUENCE SAMPLING

AC013381

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (09-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 166341)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castile, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeAtrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Liu, G., Locke, K., MacDonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, B., Olivari, F. M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6910757.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L3769
 Center clone name: 353_M_23

TITLE
JOURNAL

COMMENT

* NOTE: This record contains 174 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

* 1
 * 845 944: gap of 844 bp in length
 * 945 1799: contig of 855 bp in length
 * 1800 1899: gap of 100 bp
 * 1900 2775: contig of 876 bp in length
 * 2776 2875: gap of 100 bp
 * 2876 3771: contig of 896 bp in length
 * 3772 3871: gap of 100 bp
 * 3872 4761: contig of 890 bp in length
 * 4762 4861: gap of 100 bp
 * 4862 5736: contig of 875 bp in length
 * 5737 5836: gap of 100 bp
 * 5837 6640: contig of 804 bp in length
 * 6641 6740: gap of 100 bp
 * 6741 7617: contig of 877 bp in length
 * 7618 7717: gap of 100 bp
 * 7718 8596: contig of 879 bp in length
 * 8597 8696: gap of 100 bp
 * 8697 9555: contig of 859 bp in length

* 9556 9655: gap of 100 bp
 * 9656 10517: contig of 862 bp in length
 * 10518 10617: gap of 100 bp
 * 10618 11478: contig of 861 bp in length
 * 11479 11578: gap of 100 bp
 * 11579 12420: contig of 842 bp in length
 * 12421 12520: gap of 100 bp
 * 12521 13422: contig of 902 bp in length
 * 13423 13522: gap of 100 bp
 * 13523 14382: contig of 860 bp in length
 * 14383 14482: gap of 100 bp
 * 14483 15347: contig of 865 bp in length
 * 15348 15447: gap of 100 bp
 * 15448 16303: contig of 856 bp in length
 * 16304 16403: gap of 100 bp
 * 16404 17269: contig of 866 bp in length
 * 17270 17369: gap of 100 bp
 * 17370 18239: contig of 870 bp in length
 * 18240 18339: gap of 100 bp
 * 18340 19192: contig of 853 bp in length
 * 19193 19292: gap of 100 bp
 * 19293 20160: contig of 868 bp in length
 * 20161 20260: gap of 100 bp
 * 20261 21120: contig of 860 bp in length
 * 21121 21220: gap of 100 bp
 * 21221 22076: contig of 856 bp in length
 * 22077 22176: gap of 100 bp
 * 22177 23040: contig of 864 bp in length
 * 23041 23140: gap of 100 bp
 * 23141 24000: contig of 860 bp in length
 * 24001 24100: gap of 100 bp
 * 24101 24966: contig of 866 bp in length
 * 24967 25066: gap of 100 bp
 * 25067 25954: contig of 888 bp in length
 * 25955 26054: gap of 100 bp
 * 26055 26916: contig of 862 bp in length
 * 26917 27016: gap of 100 bp
 * 27017 27891: contig of 875 bp in length
 * 27892 27991: gap of 100 bp
 * 27992 28859: contig of 868 bp in length
 * 28860 28959: gap of 100 bp
 * 28960 29827: contig of 868 bp in length
 * 29828 29927: gap of 100 bp
 * 29928 30828: contig of 901 bp in length
 * 30829 30928: gap of 100 bp
 * 30929 31742: contig of 814 bp in length
 * 31743 31842: gap of 100 bp
 * 31843 32699: contig of 857 bp in length
 * 32700 32799: gap of 100 bp
 * 32800 33662: contig of 863 bp in length
 * 33663 33762: gap of 100 bp
 * 33763 34638: contig of 876 bp in length
 * 34639 34738: gap of 100 bp
 * 34739 35603: contig of 865 bp in length
 * 35604 35703: gap of 100 bp
 * 35704 36578: contig of 875 bp in length
 * 36579 36678: gap of 100 bp
 * 36679 37519: contig of 841 bp in length
 * 37520 37619: gap of 100 bp
 * 37620 38483: contig of 864 bp in length
 * 38484 38583: gap of 100 bp
 * 38584 39427: contig of 844 bp in length
 * 39428 39527: gap of 100 bp
 * 39528 40387: contig of 860 bp in length
 * 40388 40487: gap of 100 bp
 * 40488 41342: contig of 855 bp in length
 * 41343 41442: gap of 100 bp
 * 41443 42291: contig of 849 bp in length
 * 42292 42391: gap of 100 bp
 * 42392 43246: contig of 855 bp in length
 * 43247 43346: gap of 100 bp
 * 43347 44206: contig of 860 bp in length
 * 44207 44306: gap of 100 bp

		44307 45149: contig of 843 bp in length	
*	45150	45249: gap of 100 bp	
*	45250	46137: contig of 888 bp in length	
*	46138	46237: gap of 100 bp	
*	46238	47122: contig of 885 bp in length	
*	47123	47222: gap of 100 bp	
*	47223	48148: contig of 926 bp in length	
*	48149	48248: gap of 100 bp	
*	48249	49112: contig of 864 bp in length	
*	49113	49212: gap of 100 bp	
*	49213	50075: contig of 863 bp in length	
*	50076	50175: gap of 100 bp	
*	50176	51027: contig of 852 bp in length	
*	51028	51127: gap of 100 bp	
*	51128	51982: contig of 855 bp in length	
*	51983	52082: gap of 100 bp	
*	52083	52941: contig of 859 bp in length	
*	52942	53041: gap of 100 bp	
*	53042	53907: contig of 866 bp in length	
*	53908	54007: gap of 100 bp	
*	54008	54858: contig of 851 bp in length	
*	54859	54958: gap of 100 bp	
*	54959	55820: contig of 862 bp in length	
*	55821	55920: gap of 100 bp	
*	55921	56793: contig of 873 bp in length	
*	56794	56893: gap of 100 bp	
*	56894	57786: contig of 893 bp in length	
		36.8%; Score 571.6; DB 2; Length 166341;	
Query Match			
Best Local Similarity 91.1%; Pred. No. 1.7e-167;			
Matches 686; Conservative 0; Mismatches 55; Indels 12; Gaps 7;			
QY	774	AGCAGCCTGTGGCTTCAACCCCCACACAGTGTGACATGGGGGAAACAAAGAGATAAT	833
Db	132003	AGCAGCCTGTGGCTTCAA-CCATTACCACAGTGTGACATGGGGGAAACACGAA-GATAAAT	132060
QY	834	AGATCCTCAGTTCGACGTCTCTCCCTCAATCCAGCTCTTCCATACCCAC	893
Db	132061	AGATCCTCAGTTCGACGTCTCTCCCTCAATCCAGCTCTTCCATACCCAC	132120
QY	894	AAGCCTGAGGCTGACGCTTTTATTTATGTTTCCCTTTGGCTGTGACTGGTGGGCGAG	953
Db	132121	AAGCCTGAGGCTGACGCTTTTATTTATGTTTCCCTTTGGCTGTGACTGGTGGGCGAG	132180
QY	954	CATGACGCTTCGATTTTAAAGAGGATCTAGGGAATGTGACGACCCCTACAGGAAGGC	1013
Db	132181	CATGACGCTTCGATTTTAAAGAGGATCTAGGGAATGTGACGACCCCTACAGGAAGGC	132240
QY	1014	CTGCCATGCTGGCCCACTGTTTCACTGGAGCAAGAAAGAGATCTCATAGGACGGAGGG	1073
Db	132241	CTGCCATGCTGGCCCACTGTTTCACTGGAGCAAGAAAGAGATCTCATAGGACGGAGGG	132300
QY	1074	GGAAATGGTTTCCTCCCAAGCTGGGTGAGTGTGTTAACTGCTTATCAGCTATTTCAGACA	1133
Db	132301	GGAAATGGTTTCCTCCCAAGCTGGGTGAGTGTGTTAACTGCTTATCAGCTATTTCAGACA	132360
QY	1134	TCTCCATGTTTCTCCATGAACTCTGTGGTTTCATCATTCCTTCTTAGTTGACCTGCAC	1193
Db	132361	TCTCCATGTTTCTCCATGAACTCTGTGGTTTCATCATTCCTTCTTAGTTGACCTGCAC	132420
QY	1194	ACGTTGGTTAGACCTAGATTTAACCCCTAGGTAAGTGTGGGGTATAGAACGCTAAGAA	1253
Db	132421	ACGTTGGTTAGACCTAGATTTAACCCCTAGGTAAGTGTGGGGTATAGAACGCTA-AA	132479
QY	1254	TTTTCCCCCAAGGACCTTGTCTTCCCTTAAGCCCTTCTGGCTTCGTTTATGGTCTTCATTA	1313
Db	132480	ATTTCCCCCAAGGACCTTGTCTTCCCTTAAGCCCTTCTGGCTTCGTTTATGGGCTTCATTA	132539
QY	1314	AAAGTATAAGCCCTAATTTTGTGCTAGTCTCTAAGGAGAAACCTTTAAACCAAAAG-TTTT	1372
Db	132540	AAAGTATAAACCTAATTTTGTGCTAGTCTCTAAGGAGAAACCTTTAAACCAAAAGTTTTT	132599
QY	1373	TATCATTTAGACAAATATTGAACAACCCCC- --TATTTTGTGGGATTGAGAAGGGGTGA	1429

Db	132600	TATATTGAAAACAATATTGAACAACCCCTATTTTTTTGGGGAATTGAAAAGGGTGAA	132659
QY	1430	ATAGAGGCTT---GAGACTTTCTCTTGTGTAGGACTTTGGAGGAGAAA--TCCCCTGG	1484
Db	132660	ATAAAGACTTGAAAATTTTCTCTTTTGTGTAGGACTTTGAAGAAAAATCCCTTGG	132719
QY	1485	ACTTTCACCTAACCTCTGACATACTCCCCACAC	1517
Db	132720	ACTTTAACTAAACCTTTGGACATATCTCCCCAC	132752

Search completed: July 12, 2003, 10:39:14
Job time : 2798 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:52:55 ; Search time 40 Seconds
(without alignments)
620.067 Million cell updates/sec

Title: US-09-954-846-2

Perfect score: 1359

Sequence: 1 MAVLAPLIALVYSPRLSRW.....EOPVASTPTTVSDGENKKDK 258

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR-73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	391	28.8	265	2	S72570
2	173	12.7	253	2	T48523
3	125	9.2	135	2	C90393
4	124	9.1	104	1	TXBY1
5	122.5	9.0	117	2	E70107
6	122	9.0	103	1	TXBY2
7	121.5	8.9	145	2	AD0397
8	121	8.9	664	2	S44756
9	120.5	8.9	105	2	H69517
10	120.5	8.9	134	2	C69410
11	119	8.8	491	2	S71862
12	119	8.8	493	2	T34092
13	117.5	8.6	115	2	T29044
14	116	8.5	250	2	E87921
15	116	8.5	606	2	T31557
16	115.5	8.5	359	2	T03644
17	114	8.4	102	2	S73896
18	114	8.4	139	2	G70464
19	112.5	8.3	359	2	T37880
20	112	8.2	105	2	AG2042
21	111.5	8.2	127	2	B91218
22	111.5	8.2	127	2	C86064
23	111	8.2	643	1	S32476
24	110.5	8.1	104	2	S77780
25	110	8.1	108	1	S02802
26	110	8.1	122	2	T04090
27	110	8.1	364	2	T09614
28	110	8.1	410	2	T25574
29	109	8.0	584	2	S06318

30 109 8.0 638 1 ISMSER protein disulfide-
31 108.5 8.0 105 2 S76386 thiorodoxin M-1 -
32 108.5 8.0 139 2 E65036 rotable thiorodoxi
33 108.5 8.0 139 2 G85904 probable thiorodox
34 108.5 8.0 139 2 H91059 probable thiorodox
35 108.5 8.0 645 1 A23723 protein disulfide-
36 107.5 7.9 144 2 B82423 thiorodoxin 2 VCAO
37 107 7.9 1023 2 G96509 protein F27F5.21
38 106.5 7.8 443 2 D86183 hypothetical prote
39 106.5 7.8 496 2 A54757 protein disulfide-
40 106 7.8 102 1 G64213 thiorodoxin - Myco
41 106 7.8 102 2 D72052 thiorodoxin CP0088
42 106 7.8 102 2 H86572 thiorodoxin [impor
43 106 7.8 116 2 B70851 probable trxC prot
44 105.5 7.8 126 1 S16590 thiorodoxin h1 - c
45 105.5 7.8 788 2 T26967 hypothetical prote

ALIGNMENTS

RESULT 1

S72570

hypothetical protein C35D10.1n - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 31-Oct-1997

C:Accession: S72570

R:Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burto
ulton, L.; Gardner, A.; Green, P.; Hawkins, T.; Hillier, L.; Jier, M.; Johnston, L.;
B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D.
submitted to the EMBL Data Library, February 1995

A:Authors: Shownkeen, R.; Smaldon, N.; Smith, A.; Sonhammer, E.; Staden, R.; Sulston

proat, J.; Wohldman, P.

A:Description: The C. elegans genome project: Contiguous nucleotide sequence of over

A:Reference number: S72566

A:Accession: S72570

A:Molecule type: DNA

A:Residues: 1-265 <WIL>

A:Cross-references: EMBL:U21324; NID:G687879; PID:G687889

A:Experimental source: strain Bristol N2

C:Genetics:

A:Map position: 3

A:Introns: 52/3; 111/1; 213/2

A:Note: C35D10.10

Query Match 28.8%; Score 391; DB 2; Length 265;

Best Local Similarity 32.2%; Pred. No. 5.4e-28;

Matches 84; Conservative 52; Mismatches 81; Indels 44; Gaps 6;

QY 14 VPRLS--RWLAQPYLLSALLSAFLVLRKLPPLCHGLPTOREDCNCFDHWREVEILMF 71

Db 3 IPRLDEVRRALTAFHFFNTLLALAPPVIRS-TSLGDYV-FAVEGNEQCEIDSREREILMF 60

QY 72 LSAIVMMKNRRS-----MELMTC-----KPL 93

Db 61 LLILAWKGRATNMVHMVNNIFLFSKIAGMFLFTRADILPGIITLACILVTLVLPPEV 120

QY 94 YMGPEYIKYFNDKTIDEELERDKRYTWIWEFFANMNDQSFAPYADLSKYNCTGNF 153

Db 121 YNGPEQVTFQGEQLFEELTRNRTIWIQFTTWSPECRHTSPVFAELSKQFTLPNKKF 180

QY 154 GKVDVGRYDVTYRYKYSTSLTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTSEENV 213

Db 181 KGLDIGRWAKGEGRFVNAHPMSRQLPTICVFKDAKEIARRPLVNDSSRAVPFVFESENC 240

QY 214 IREENLNELY--QRAKKLSKA 232

Db 241 VLAFDLLNLYNEQKEKKGAKA 261

RESULT 2

T48523

hypothetical protein T22p22.30 - Arabidopsis thaliana

[illegible]

A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AD0397
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-145 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92504.1; PID:g15981204; GSPDB:GN00175
C:Genetics:
A:Gene: trxC
C:Superfamily: thioredoxin; thioredoxin homology
Query Match 8.9%; Score 121.5; DB 2; Length 145;
Best Local Similarity 27.7%; Pred. No. 0.00088;
Matches 33; Conservative 30; Mismatches 33; Indels 23; Gaps 6

hypothetical protein F56G4.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence revision 29-Oct-1999 #t31557
RESULT 15

Search completed: July 9, 2003, 11:58:00
Job time : 41 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 11:39:55 ; Search time 11 seconds
(without alignments)
972.808 Million cell updates/sec

Title: US-09-954-846-2
Perfect score: 1359
Sequence: 1 MAVLAPLIALVSVPLRSRW.....EQPVASTPTVSDGENKKDK 258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	9.1	103	1	TRX1_YEAST
2	122.5	9.0	107	1	THIO_BORBU
3	122	9.0	102	1	TRX2_YEAST
4	121	8.9	618	1	PD44_CAEEL
5	117.5	8.6	115	1	TH11_CAEEL
6	116	8.5	127	1	TH10_NEUCR
7	114	8.4	102	1	THIO_MYCPM
8	113	8.3	106	1	PD12_SCHPO
9	112.5	8.3	359	1	PD44_RAT
10	111	8.2	643	1	PD12_DROME
11	110.5	8.1	496	1	PD12_CORNE
12	110	8.1	107	1	TH12_ORYSA
13	110	8.1	122	1	TH12_ORYSA
14	110	8.1	364	1	PD46_MEDSA
15	109	8.0	638	1	PD44_MOUSE
16	108.5	8.0	105	1	TH11_SYNV3
17	108.5	8.0	139	1	TH12_ECOLI
18	108.5	8.0	359	1	TIGA_ASPNG
19	108.5	8.0	645	1	PD44_HUMAN
20	108	7.9	102	1	TH10_CHLPS
21	106	7.8	102	1	TH10_CHLPS
22	106	7.8	102	1	TH10_MYCCE
23	106	7.8	108	1	TH10_BUCAP
24	106	7.8	115	1	TH10_MYCCT
25	105.5	7.8	126	1	TH11_TOBAC
26	105	7.7	509	1	PD11_RABIT
27	105	7.7	510	1	PD11_BOVIN
28	104.5	7.7	119	1	TH22_BRANA
29	104.5	7.7	458	1	TRXB_MYCLE
30	104.5	7.7	522	1	PD11_YEAST
31	104	7.7	108	1	TH10_BUCAP
32	102.5	7.5	105	1	TH10_TREPA
33	102.5	7.5	118	1	TH55_ARATH

34	102.5	7.5	497	1	BS2_TRYBB	P12865 trypanosoma
35	102	7.5	106	1	TH11_ANASP	P06544 anabaena sp
36	102	7.5	133	1	TH22_ARATH	O38879 arabidopsis
37	102	7.5	361	1	PD46_ARATH	O22263 arabidopsis
38	102	7.5	369	1	ER38_NEUCR	O92249 neurospora
39	102	7.5	509	1	PD11_RAT	P04785 rattus norv
40	101.5	7.5	119	1	TH44_ARATH	O39239 arabidopsis
41	101	7.4	509	1	PD11_MOUSE	P09103 mus musculu
42	100	7.4	104	1	TH13_DICDI	P29447 dictyostell
43	100	7.4	106	1	TH10_PENCH	P34723 penicillium
44	99.5	7.3	118	1	TH33_ARATH	O42403 arabidopsis
45	99.5	7.3	175	1	TH11_WHEAT	O92p21 triticum ae

ALIGNMENTS

RESULT 1					
TRX1_YEAST					
ID	TRX1_YEAST	STANDARD;	PRT;	103 AA.	
AC	P22803;				
DT	01-AUG-1991 (Rel. 19, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Thioredoxin I (TR-I)				
GN	TRX1 OR YGR209C OR G7746.				
OS	Saccharomyces cerevisiae (Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.				
OX	NCBI_TaxID=4932;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=91107668; PubMed=1988444;				
RA	Gan Z.-R.;				
RT	"yeast thioredoxin genes."				
RL	J. Biol. Chem. 266:1692-1696(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=91225027; PubMed=2026619;				
RA	Muller E.G.D.;				
RT	"Thioredoxin deficiency in yeast prolongs S phase and shortens the G1 interval of the cell cycle."				
RL	J. Biol. Chem. 266:9194-9202(1991).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	STRAIN=S288c;				
RA	MEDLINE=97060019; PubMed=8904340;				
RT	Guereiro P., Barreiros T., Soares H., Cyrne L., Maia e Silva A., Rodrigues-Pousada C.;				
RL	"Sequencing of a 17.6 kb segment on the right arm of yeast chromosome VII reveals 12 ORFs, including CCT, ADE3 and TR-I genes, homologues of the yeast PMT and EFC genes, of the human and bacterial electron-transferring flavoproteins (beta-chain) and of the Escherichia coli phosphoserine phosphohydrolase, and five new ORFs."				
RL	Yeast 12:273-280(1996).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	STRAIN=GRF88;				
RA	Song J.M., Cheung E., Rabinowitz J.C.;				
RL	Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE OF 1-12.				
RX	STRAIN=YN 295;				
RA	MEDLINE=95050519; PubMed=7961686;				
RL	Chae H.Z., Chung S.J., Rhee S.G.;				
RT	"Thioredoxin-dependent peroxide reductase from yeast."				
RL	J. Biol. Chem. 269:27670-27678(1994).				
RN	[6]				
RP	SEQUENCE OF 26-42.				
RX	MEDLINE=72104597; PubMed=4333602;				
RA	Hall D.E., Baldesten A., Holmgren A., Reichard P.;				
RT	"The covalently-bound flavin of hepatic monoamine oxidase. 2. Identification and properties of cysteinyl riboflavin."				


```

RN RP SEQUENCE OF 1-12.
RC STRAIN=YN 295;
RX MEDLINE=95050519; PubMed=7961686;
RA Chee H.Z., Chung S.J., Rhee S.G.;
RT "Thioredoxin-dependent peroxide reductase from yeast.";
RL J. Biol. Chem. 269:27670-27678(1994).
RN RP SEQUENCE OF 25-33.
RX MEDLINE=72104597; PubMed=4333602;
RA Hall D.E., Baldesten A., Holmgren A., Reichard P.;
RT "The covalently-bound flavin of hepatic monoamine oxidase. 2.
RT Identification and properties of cysteinyl riboflavin.";
RL Eur. J. Biochem. 24:328-335(1971).
CC CC -!- FUNCTION: Participates in various redox reactions through the
CC reversible oxidation of its active center dithiol to a disulfide
CC and catalyzes dithiol-disulfide exchange reactions.
CC -!- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M62647; AAA35177.1; -
CC EMBL; M59169; AAA35171.1; -
CC EMBL; Z73215; CAA97572.1; -
CC PIR; S15048; TXBY2.
CC PIR; A39847; A39847.
CC PIR; B38669; B38669.
CC HSP; P10599; IERV.
CC SGD; S0003441; TRX2.
CC InterPro; IPR000063; Thioered.
CC Pfam; PF00085; Thioered; 1.
CC PRINTS; PR00421; THIREDOXIN.
CC TIGRFAMs; TIGR01068; thioredoxin; 1.
CC PROSITE; PS00194; THIREDOXIN; 1.
KW Redox-active center; Electron transport.
FT INIT_MET 0
FT DISULFID 29 32 REDOX-ACTIVE (BY SIMILARITY).
SQ SEQUENCE 102 AA; 11104 MW; 446EA348281B6C0C CRC64;

Query Match 9.08; Score 122; DB 1; Length 102;
Best Local Similarity 34.18; Pred. No. 0.00013;
Matches 29; Conservative 17; Mismatches 29; Indels 10; Gaps 3;

QY 109 DEELERDKRVTWIVFEFANKNSDCOSFAPYADLSLKYNCTGLNFGKVDVGRYTDVSTRY 168
DB 11 DSAIAQDKLV--VDFYATWCGPKMTAPMIEKFESEQY--PQADFYLKLDVDELGDVAQKN 66
QY 169 KVSPTSPLTKOLPTLLIFQGGKEAMR 193
DB 67 EVSA-----MPTLLLFKNCKEVAK 85

RESULT 4.
PDA4_CAEEL
ID PDA4_CAEEL STANDARD; PRT; 618 AA.
AC P34329;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable protein disulfide isomerase A4 precursor (EC 5.3.4.1) (Exp-72
DE homolog).
GN C14B9.2
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]

```

```

RP SEQUENCE FROM N. A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
CC CC -!- FUNCTION: NOT KNOWN.
CC -!- CATALYTIC ACTIVITY: Rearrangement of both intrachain and
CC interchain disulfide bonds in proteins to form the native
CC structures.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC -!- SIMILARITY: CONTAINS 3 THIREDOXIN DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L15188; AAM22024.1; -
CC PIR; S44756; S44756.
CC HSP; P07237; IMEK.
CC MEROPS; C17.002; -.
CC WormPep; C14B9.2; CE00073.
CC InterPro; IPR000886; ER_target.
CC InterPro; IPR000063; Thioered.
CC Pfam; PF00085; thioered; 3.
CC PRINTS; PR00421; THIREDOXIN.
CC TIGRFAMs; TIGR01126; pdi_dom; 3.
CC TIGRFAMs; TIGR01130; ER_PDI_fam; 1.
CC PROSITE; PS00014; ER_TARGET; 1.
CC PROSITE; PS00194; THIREDOXIN; 3.
KW Hypothetical protein; Isomerase; Redox-active center;
KW Endoplasmic reticulum; Repeat; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 618 PROBABLE PROTEIN DISULFIDE ISOMERASE A4.
FT DISULFID 65 68 REDOX-ACTIVE (BY SIMILARITY).
FT DISULFID 176 179 REDOX-ACTIVE (BY SIMILARITY).
FT DISULFID 529 532 REDOX-ACTIVE (BY SIMILARITY).
FT SITE 615 618 PREVENT SECRETION FROM ER.
SQ SEQUENCE 618 AA; 69797 MW; 1DC0207A71444220 CRC64;

Query Match 8.98; Score 121; DB 1; Length 618;
Best Local Similarity 25.18; Pred. No. 0.0017;
Matches 51; Conservative 22; Mismatches 72; Indels 58; Gaps 9;

QY 48 GLPTQR--EDG-NPCDFD-----WREIILFLSAIYMKNRSMFLMTCKPLY 94
DB 105 GYPTLKWKGKGNPDYDGGDEAGIVEWE-----SRVDPNYKPP-- 145
QY 95 MGPEYIKYFNKDTIDEELERDKRVTWIVFEFANKNSDCOSFAPYADLSLKYNCTG--LN 152
DB 146 --PEVTLITENFDDFISNNELV--LVEFTAPCGHCKKLAPEYKAAQKLKAGSKVK 201
QY 153 FGKVDVGRYTDVSTRYKVSPTSPLTKOLPTLLIFQGGKEAMRRPQIDKGRAVSWTFSEN 212
DB 202 LGKVDATIEKDLGTGYGVGYPTMK-----IIRNGRRFDYNGPREA 242
QY 213 V-IREFNLNELYQRAKKLSKAGD 234

```

```
Db 243 AGIIKYMTDOSKPAKKLPKLD 265
RESULT 5
THIO_CAEEL
ID THIO_CAEEL STANDARD; PRT; 115 AA.
AC Q09433;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable thioredoxin B0228.5.
GN B0228.5.
OS Caenorhabditis elegans.
CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
CC Rhabditidae; Peleoderinae; Caenorhabditis.
CC NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Bristol N2.
RC Leimbach D., Waterston R.;
RA Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RL FUNCTION: Participates in various redox reactions through the
CC reversible oxidation of its active center dithiol to a disulfide
and catalyzes dithiol-disulfide exchange reactions (By
similarity).
CC -|- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U23168; AAC38808.1;
CC HSSP: P07237; IMEK.
CC WormPep: B0228.5; CE01745.
CC InterPro: IPR000063; ThioRed.
CC Pfam: PF00085; ThioRed; 1.
CC PRINTS: PR00421; THIOREDOXIN.
CC PROSITE: PS00194; THIOREDOXIN; 1.
CC KW Hypothetical protein; Redox-active center; Electron transport.
FT DISULFID 39 42 REDOX-ACTIVE (BY SIMILARITY).
SQ SEQUENCE 115 AA; 13323 MW; 073F51A3EA97AD4B CRC64;
Query Match 8.6%; Score 117.5; DB 1; Length 115;
Best Local Similarity 29.2%; Pred. No. 0.0004;
Matches 26; Conservative 20; Mismatches 34; Indels 9; Gaps 3;
QY 100 IKYFNKTIIDELEKRVTWIIEFFANNSDCQSFAPYADLSLKYNCTGLNFGKVDVG 159
11 VKYF-QSDFEQLIROPKRIIILDFYATWCGPKAIAPLYKELATTHK--GIIFCKVDVD 67
QY 160 RYTDVSTRYKVSPLTKOLPTLLIFOGG 188
Db 68 EAEDLCSKYDV-----KMPFTFTKNG 90
RESULT 6
THIO_NEUCR
ID THIO_NEUCR STANDARD; PRT; 127 AA.
AC P42115;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thioredoxin.
GN TRX.
OS Neurospora crassa.
CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
CC Sordariales; Sordariaceae; Neurospora.
CC NCBI_TaxID=5141;
RN [1]
SEQUENCE FROM N.A.
RP Akiyama M., Nakashima H.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: Participates in various redox reactions through the
CC reversible oxidation of its active center dithiol to a disulfide
and catalyzes dithiol-disulfide exchange reactions.
CC -|- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D45892; BAA08305.1;
CC HSSP: P10599; ITRV.
CC InterPro: IPR000063; ThioRed.
CC Pfam: PF00085; ThioRed; 1.
CC PRINTS: PR00421; THIOREDOXIN.
CC PROSITE: PS00194; THIOREDOXIN; 1.
CC KW Redox-active center; Electron transport.
FT DISULFID 33 36 REDOX-ACTIVE (BY SIMILARITY).
SQ SEQUENCE 127 AA; 13716 MW; 9F8A02C88EA492EB CRC64;
Query Match 8.5%; Score 116; DB 1; Length 127;
Best Local Similarity 25.0%; Pred. No. 0.0063;
Matches 28; Conservative 27; Mismatches 35; Indels 22; Gaps 4;
QY 121 IVEFFANNSDCQSFAPYADLSLKYNCTGLNFGKVDVGRTYSTRYKVSPLTKQL 179
Db 25 VADFADWCGPKAIAPYAFKTFSPNLFAPKINVDVSVQVQVAHYRVA-----M 78
QY 180 PTLLIFQGGKAMRRPQIDKKGRAVSWTFSEENVIREFNELNYQRAKKLSK 231
Db 79 PTELFKNGKQV-----AVNGSV---MIOGADVNSLRRAAEKNGR 115
RESULT 7
THIO_MYCPN
ID THIO_MYCPN STANDARD; PRT; 102 AA.
AC P75512;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thioredoxin (TRX).
GN TRXA OR TRX OR MPN263 OR MP570.
OS Mycoplasma pneumoniae.
CC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
CC NCBI_TaxID=2104;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Ben-Menachem G., Himmelreich R., Herrmann R., Aharonowitz Y.,
RA Rotten S.;
RT "The thioredoxin reductase system of mycoplasmas.";
RL Microbiology 143:1933-1940(1997).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -|- FUNCTION: Participates in various redox reactions through the
CC reversible oxidation of its active center dithiol to a disulfide
and catalyzes dithiol-disulfide exchange reactions (By
similarity).
CC -|- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY.
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U51987; AAC45450.1; -;
 DR EMBL: AE000056; AAB96218.1; -;
 DR HSSP: P10599; IERV.
 DR InterPro: IPR000063; ThioRed.
 DR Pfam: PF00085; thioRed; 1.
 DR PRINTS: PR00421; THIOREDOXIN.
 DR TIGRFAMs: TIGR01068; thioredoxin; 1.
 DR PROSITE: PS00194; THIOREDOXIN; 1.
 DR Redox-active center; Electron transport; Complete proteome.
 FT DISULFID 30 33 REDOX-ACTIVE (BY SIMILARITY).
 SQ SEQUENCE 102 AA; 11215 MW; 0D17B97E976FC144 CRC64;
 Query Match 8.4%; Score 114; DB 1; Length 102;
 Best Local Similarity 29.5%; Pred. No. 0.00073;
 Matches 26; Conservative 21; Mismatches 31; Indels 10; Gaps 3;
 QY 106 KTIDEELERDKRVTWIVFFFWNSDCQSFAPYADLSLKYNCTGLNFGKVDVGRYTDYS 165
 Db 9 KQELGFASNNKV--IIDFWAECPCPKITGFEFAKAASEVSTA--FAKVNVDQTDIA 64
 QY 166 TRYKVTSPKQLPTLILFOGGKEAMR 193
 Db 65 AAYKITS-----LPTIVLFEKGQEKHR 86
 RESULT 8
 THIO_COPCM STANDARD; PRT; 106 AA.
 ID Q9UW02, 2001 (Rel. 40, Created)
 AC Q9UW02, 2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Thioredoxin (Allergen Cop c 2).
 OS Coprinus comatus (Shaggy mane).
 OC Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 AG Agaricales; Agaricaceae; Coprinus.
 NCBI_TaxID=56187;
 [1]
 SEQUENCE FROM N.A.
 RA Brander K.A., Crameri R., Schuermann P., Pichler W.J., Helbling A.;
 RT "Coprinus thioredoxin as inhalative allergen and crossreactive human
 RL autoantigen".;
 CC Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Participates in various redox reactions through the
 CC reversible oxidation of its active center dithiol to a disulfide
 CC and catalyzes dithiol-disulfide exchange reactions.
 CC -!- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AJ242791; CAB52130.1; -;
 DR HSSP: P80028; ITOF.
 DR InterPro: IPR000063; ThioRed.
 DR Pfam: PF00085; thioRed; 1.
 DR PRINTS: PR00421; THIOREDOXIN.
 DR TIGRFAMs: TIGR01068; thioredoxin; 1.
 DR PROSITE: PS00194; THIOREDOXIN; 1.
 DR Redox-active center; Electron transport; Allergen.
 KW

FT DISULFID 30 33 REDOX-ACTIVE (BY SIMILARITY).
 SQ SEQUENCE 106 AA; 11772 MW; 05A2155B210E8C69 CRC64;
 Query Match 8.3%; Score 113; DB 1; Length 106;
 Best Local Similarity 30.0%; Pred. No. 0.00095;
 Matches 21; Conservative 17; Mismatches 26; Indels 6; Gaps 1;
 QY 121 IVEFFANWNSDCQSFAPYADLSLKYNCTGLNFGKVDVGRYTDVSTYKYSTPLTKQLP 180
 Db 22 IIDFWATWCGPCRCVPIFEKSEKYGANNVFAKVVDVTASDISEAKI-----RAMP 75
 QY 181 TLILFOGGKE 190
 Db 76 TFOVYKGGQK 85
 RESULT 9
 PD12_SCHPO STANDARD; PRT; 359 AA.
 ID PD12_SCHPO
 AC Q13811, 1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative protein disulfide isomerase C17H9.14c precursor (EC 5.3.4.1).
 GN SPAC17H9.14c.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OC NCBI_TaxID=4896;
 [1]
 SEQUENCE FROM N.A.
 RP STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprez B.,
 RA Weltyens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Borzym K., Langer I., Beck A., Leirach H., Wambutt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Leirach H., Wambutt R., Pohl T.M.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe".;
 CC Nature 415:871-880(2002).
 CC -!- FUNCTION: PARTICIPATES IN THE FOLDING OF PROTEINS CONTAINING
 CC DISULFIDE BONDS, MAY BE INVOLVED IN GLYCOSYLATION, PROLYL
 CC HYDROXYLATION AND TRIGLYCERIDE TRANSFER (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: Rearrangement of both intrachain and
 CC interchain disulfide bonds in proteins to form the native
 CC structures.
 CC -!- SIMILARITY: CONTAINS 2 THIOREDOXIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial


```

Db      13 NKDEFDAQMTKAEAGAKGWIIDFTASWCGPCRFAPVFAEYAKKF--PGAFLVKVDVDEL 70
QY      162 TDYSTRYKVSTSLTKOLPTLILFQGKEA 191
           :|::|| ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      71 KEVAEKYNV-----EAMPTFLFIKDGAEA 94

RESULT 14
PDA6_MEDSA          STANDARD;             PRT:   364 AA.
AC    P38661;
DT     01-FEB-1995 (Rel. 31, Created)
DT     01-FEB-1995 (Rel. 31, Last sequence update)
DT     01-FEB-1995 (Rel. 31, Last annotation update)
DE     Probable protein disulfide isomerase A6 precursor (EC 5.3.4.1) (P5).
OS     Medicago sativa (Alfalfa).
OC     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC     Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC     eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX     NCBI_TaxID=3879;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93272046; PubMed=1303795;
RA      Shorrosh B.S.; Dixon R.A.;
RT      "Molecular characterization and expression of an alfalfa protein with
RT      sequence similarity to mammalian ERP72, a glucose-regulated
RT      endoplasmic reticulum protein containing active site sequences of
RT      protein disulphide isomerase.";
RL      Plant J. 2:51-58(1992).
RC     -I- CATALYTIC ACTIVITY: Rearrangement of both intrachain and
CC       interchain disulfide bonds in proteins to form the native
CC       structures.
CC     -I- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC     -I- SIMILARITY: CONTAINS 2 THIOREDUXIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M80235; AAB46930.1; -.
CC DR      HSPS; P07237; IMEX.
CC DR      InterPro: IPRO00063; Thioired.
CC DR      Pfam; PF00085; thioired; 2.
CC DR      PRINTS; PR00421; THIOREDUXIN.
CC DR      TIGRPFAMS; TIGR01126; pdi_dom; 2.
CC DR      PROSITE; PS00194; THIOREDUXIN; 2.
CC KW      Redox-active center; Isomerase; Endoplasmic reticulum; Repeat; Signal.
CC FT      SIGNAL        1      28      POTENTIAL.
CC FT      CHAIN         29      364      PROBABLE PROTEIN DISULFIDE ISOMERASE A6.
CC FT      DISULFID      58      61      REDOX-ACTIVE (BY SIMILARITY).
CC FT      DISULFID     177     180      REDOX-ACTIVE (BY SIMILARITY).
CC SQ      SEQUENCE     364 AA; 40492 MW; 7496F4093C3A01E6 CRC64;

Query Match              8.1%; Score 110; DB 1; Length 364;
Best Local Similarity    25.6%; Pred. No. 0.0089;
Matches    40; Conservative    24; Mismatches    60; Indels    32; Gaps    8

QY      97 PEIKYFNDKTIDE-ELERDKRWTVIEFFANWSNDQCQSFAPIYADLSLKYNC-TGLNFG 154
           |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      146 PSHVVLTPTTFNEVLDTGTDV--LVFEYPWCCHCKSLAPIYEKVAVPKSEDDVIA 203

QY      155 KVDVGRYDVSTRYKVSTSLTKOLP-----TILFGKGKAMRRPQID 198
           ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      204 NLDAQRYDLAEKYDVSGFTPLFPFKGNKAGEDYGGGRDLLDDVEAFINEKSGTSR---D 260

QY      199 KKGRAVSWTFSEENVIREFNLNELIQRAKLISKAGD 234
           ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      261 AKQLGT----SEAGIVED--LDEL-----VKBEFVAAND 287

```

```
RESULT 15
PDA4_MOUSE
ID PDA4_MOUSE STANDARD; PRT: 638 AA.
AC P08003; P15841;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein disulfide isomerase A4 precursor (EC 5.3.4.1) (Protein Exp-72)
DE (Exp72).
GN PDIA4 OR ERP72 OR CAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=90110091; PubMed=2295602;
Mazzarella R.A., Srinivasan M., Haugjorden S.M., Green M.;
"Erp72, an abundant luminal endoplasmic reticulum protein, contains
three copies of the active site sequences of protein disulfide
isomerase.";
RL J. Biol. Chem. 265:1094-1101(1990).
RN [2]
SEQUENCE OF 35-638 FROM N.A.
RC STRAIN-BALB/C; TISSUE-T-cell;
RX MEDLINE=88096388; PubMed=3320970;
RA Gough N.M., King J.A., Dunn A.R.;
"An mRNA sharing sequences with a variant granulocyte-macrophage
colony stimulating factor cDNA clone.";
RL Nucleic Acids Res. 15:10584-10584(1987).
CC -!- CATALYTIC ACTIVITY: Rearrangement of both intrachain and
interchain disulfide bonds in proteins to form the native
structures.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -!- SIMILARITY: CONTAINS 3 THIOREDOXIN DOMAINS.
CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 544
ONWARD DUE TO A FRAMESHIFT.
CC -----
This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
EMBL; J05186; AAA39907.1; -
DR EMBL; Y00884; CAA68777.1; ALT_FRAME.
DR PIR; B34930; ISMSER.
DR PIR; S06318; S06318.
DR HSSP; P07237; IMEK.
DR MEROPS; C17.002; -.
DR MGD; MGI:104864; Cal.
DR InterPro; IPR000886; ER_target.
DR InterPro; IPR000063; ThioRed.
DR Pfam; PF00085; thioRed; 3.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMs; TIGR01126; pdi_dom; 3.
DR TIGRFAMs; TIGR01130; ER_PDI_fam; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00194; THIOREDOXIN; 3.
KW Redox-active center; Isomerase; Endoplasmic reticulum; Repeat; Signal.
FT SIGNAL 1 20
FT CHAIN 21 638
FT CARBOHYD 36 36
FT DOMAIN 39 50
FT DISULFID 84 87
FT DISULFID 199 202
FT DISULFID 548 551
FT SITE 635 638
FT SITE 638 AA: 71973 MW; FA06D250D06A27BB CRC64;
SQ SEQUENCE
```

```
Query Match 8.0%; Score 109; DB 1; Length 638;
Best Local Similarity 24.2%; Pred. No. 0.023;
Matches 38; Conservative 27; Mismatches 60; Indels 32; Gaps 5;
OY 104 NDKTIDEEL-ERDKRVTWIVEFFANWNCOSFAPYIADL--SLKYNCTGLNFGKVDVGR 160
|| | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 61 NDGNFDFVADKD---TVLLEFYAPMGCHCKQFAPEYEKIASTLKDNDPPIAVAKIDATS 117
OY 161 YTDVSTRYKYSTSPLTKQLPTLLLFQGGKAMRRPQIDKKGRAVSWTFSEENVIREENLN 220
|| | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 118 ASMLASKFDVSGYPTTKIL-----KRGQAVDYDGSRTQ-----E 151
OY 221 ELYQRAKKLSKAGDNIPEEQPVASTPTVTSVDGENKKD 257
|| | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 152 EIVAKREVNSQPDWTPPEVTVLSLTKNDFDDVNNAD 188
```

Search completed: July 9, 2003, 11:55:46
Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 09:47:42 ; Search time 185 Seconds
(without alignments)
3140.625 Million cell updates/sec

Title: US-09-954-846-2

Perfect score: 1359

Sequence: 1 MAVLAPLIAIVSVPLSRW.....EQPVASTPTTVSDGKNKKDK 258

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

arched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09954846/runat_09072003_102144_12148/app_query.fasta_1.455
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09954846 -CGN_1_1_200/runat_09072003_102144_12148 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1359	100.0	774	21	AAA15981	Human protein clon
2	1359	100.0	1527	21	AAA15991	Human protein clon
3	1359	100.0	1598	21	AAF16316	Human prostate can
4	1326	97.6	1618	20	AAX52251	Protein PRO270 CDN
5	1326	97.6	1620	22	AAS45940	Human CGI encoding
6	1326	97.6	1730	22	ABA09248	Human GGI-31 prote
7	1326	97.3	1699	19	AAV34294	Human secreted pro
8	1313	96.6	1621	22	AAF72409	Human PRO270 CDNA
9	1313	96.6	1652	19	AAV34316	Human secreted pro
10	1308	96.2	1659	24	ABL49614	Prostaglandin E2 (
11	1299	95.6	1631	22	AAF93762	Human CDNA encodin
12	984	73.1	774	21	AZ442527	Human 5' EST isola
13	956	70.3	726	22	AAF93955	Primer specific fo
14	925.5	68.1	2030	23	AAS83076	DNA encoding novel
15	622	45.8	439	21	AAA41964	Human secreted exp
16	620	45.6	461	21	AAO10008	Human secreted pro
17	579.5	42.6	3369	22	ABA15532	Human nervous syst
18	579.5	42.6	3369	22	ABA15533	Human nervous syst
19	514	37.8	444	20	AA41205	Human secreted pro
20	498	36.6	2297	22	ABA15531	Human nervous syst
21	476.5	35.1	919	23	ABL17145	Drosophila melanog
22	476.5	35.1	2919	23	ABL17145	Drosophila melanog
23	447	32.9	430	20	AA41202	Human secreted pro
24	431.5	31.8	5423	23	AAS83062	DNA encoding novel
25	431.5	31.8	5423	24	ABL68483	Kidney cancer rela
26	430	31.6	452	20	AA41203	Human secreted pro
27	424	31.2	307	23	AAS83053	DNA encoding novel
28	422	31.1	469	21	AZ433018	Human 5' EST isola
29	415	30.5	307	24	ABK45061	CDNA encoding colo
30	372	27.4	274	21	AAA45185	Human secreted exp
31	353.5	26.0	756	23	ABV21932	Human prostate exp
32	353.5	26.0	756	23	ABV27764	Human prostate exp
33	336	24.7	1895	23	AAS83073	DNA encoding novel
34	323	23.8	1409	22	ABA15530	Human nervous syst
35	280	20.6	531	21	AZ433016	Human 5' EST isola
36	269	19.8	153	22	AA119490	Human breast cance
37	269	19.8	433	22	AA109103	Human prostate exp
38	232	17.1	370	23	ABV02989	Human nervous syst
39	175	12.9	367	22	AB11541	Human prostate exp
40	163.5	12.0	375	23	ABV33303	Human prostate exp
41	163.5	12.0	375	23	ABV42226	Human prostate exp
42	163.5	12.0	389	23	ABV12158	Human prostate exp
43	130.5	9.6	3550	9	AA81540	Sequence of pTRX-2
44	130.5	9.6	3550	13	AAQ31623	pTRX-2 containing
45	129.5	9.5	3552	11	AAQ05875	pTRX-2 plasmid seq

ALIGNMENTS

RESULT 1

AAAL5981

ID AAAL5981 standard; CDNA; 774 BP.

XX AAAL5981;

AC AAAL5981;

DT 12-JUN-2000 (first entry)

DE Human protein clone HP10392 coding sequence.

XX

Human protein; hydrophobic domain; nutritional source; haematopoiesis;
cytokine production; cell proliferation; cell differentiation;
immune deficiency; infectious disease; autoimmune disorder; asthma;
multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
allergic reaction; osteoporosis; osteoarthritis; periodontal disease;
nervous system disorder; Alzheimer's disease; Parkinson's disease;
Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;
systemic cytokine damage; tissue differentiation; contraceptive; stroke;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OR nucleic - nucleic search, using sw model

Run on: July 12, 2003, 08:31:01 ; Search time 269 Seconds
(without alignments)
13018.057 Million cell updates/sec

Title: US-09-954-846-4

Perfect score: 1555

Sequence: 1 AGGGGAGCGGGCGGAGACC.....TAATAAAAGATTCGGATTA 1555

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by change to have a score greater than or equal to the score of the results being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1520	97.7	1527	21	Human protein clon
2	1511	97.2	1598	21	Human prostate can
3	1429	91.9	1730	22	Human CGI-31 prote
4	1381.6	88.8	1699	19	Human secreted pro
5	1370.2	88.1	1631	22	Human cDNA encoding
6	1370	88.1	1620	22	Human DNA encoding
7	1368	88.0	1618	20	Protein PRO270 cDN
8	1361.8	87.6	1652	19	Human secreted pro
9	1361.2	87.5	1659	24	Prostaglandin E2 (

10	1359	87.4	1621	22	AAF72409	Human PRO270 cDNA.
11	868	55.8	2297	22	ABAI5531	Human nervous syst
12	868	55.8	3369	22	ABAI5532	Human nervous syst
13	868	55.8	3369	22	ABAI5533	Human nervous syst
14	774	49.8	774	21	AAAI5981	Human protein clon
15	613	39.4	2030	23	AAS83076	DNA encoding novel
16	521	33.5	774	21	AAS43527	Human 5' EST isola
17	463	29.8	726	22	AAF93955	Primer specific fo
18	441.2	28.4	521	16	AAF22413	Human gene signatu
19	348.4	22.4	439	21	AAA41964	Human secreted exp
20	308	19.8	469	21	AAZ43018	Human 5' EST isola
21	305.6	19.7	452	20	AAZ41203	Human secreted pro
22	305.2	19.6	461	21	AAC01008	Human secreted pro
23	287.2	18.5	430	20	AAZ41202	Human secreted pro
24	261	16.8	307	23	AAS83053	DNA encoding novel
25	261	16.8	444	20	AAZ41205	Human secreted pro
26	258.4	16.6	307	24	ABK45061	CDNA encoding colo
27	252.2	16.2	5423	23	AAS83062	DNA encoding novel
28	252.2	16.2	5423	24	ABK45062	Kidney cancer rela
29	246	15.8	1895	23	AAS83073	DNA encoding novel
30	230.6	14.8	531	21	AAZ43016	Human 5' EST isola
31	221	14.2	756	23	ABV21932	Human prostate exp
32	221	14.2	756	23	ABV21932	Human prostate exp
33	200.4	12.9	274	21	AAA45185	Human secreted exp
34	144.6	9.3	153	22	AAI19490	Human breast cance
35	144.6	9.3	433	22	AAI09103	Human breast cance
36	141.8	9.1	370	23	ABV02989	Human prostate exp
37	110.2	7.1	1409	22	ABAI5530	Human nervous syst
38	95.8	6.2	2919	23	ABLI1744	Protophila melanog
39	89.2	5.9	919	23	ABLI1745	Drosophila melanog
40	82.2	5.7	367	22	ABAI1541	Human nervous syst
41	84.4	5.4	375	23	ABV33303	Human prostate exp
42	84.4	5.4	375	23	ABV33303	Human prostate exp
43	84.4	5.4	389	23	ABV12158	Human prostate exp
44	50	3.2	50	20	AAZ52405	Probe used to iso
45	50	3.2	50	22	AAF72563	Human PRO polypept

ALIGNMENTS

RESULT 1

AAAI5991

ID AAAI5991 standard; cDNA; 1527 BP.

XX AAAI5991;

AC AAAI5991;

XX 12-JUN-2000 (first entry)

DT Human protein clone HP10392 full length coding sequence.

DE Human protein; hydrophobic domain; nutritional source; haematopoiesis;

XX Human protein; hydrophobic domain; nutritional source; haematopoiesis;

KW cytokine production; cell proliferation; cell differentiation;

KW immune deficiency; infectious disease; autoimmune disorder; asthma;

KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;

KW allergic reaction; osteoporosis; osteoarthritis; periodontal disease;

KW nervous system disorder; Alzheimer's disease; Parkinson's disease;

KW Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;

KW systemic cytokine damage; tissue differentiation; contraceptive; stroke;

KW coagulation disorder; myocardial infarction; inflammatory condition;

KW septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;

KW nephritis; therapy; ss.

XX Homo sapiens.

OS Homo sapiens.

XX WO200005367-A2.

PN WO200005367-A2.

XX PD 03-FEB-2000.

XX 22-JUL-1999; 99WO-JP03929.

XX 24-JUL-1998; 98JP-0208820.

PR 07-AUG-1998; 98JP-0224105.

PR 25-AUG-1998; 98JP-0238116.
PR 09-SEP-1998; 98JP-0254736.
PR 29-SEP-1998; 98JP-0275505.
XX
FA (SAGA) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX
XX Kato S, Kimura T;
PI
XX WPI; 2000-182694/16.
DR P-PSDB; RAY94893.
XX
XX Novel human proteins having hydrophobic domains useful for treating
PT osteoporosis, Alzheimer's disease, Parkinson's disease, asthma,
PT multiple sclerosis, rheumatoid arthritis, cancer, anaemia, and stroke -
XX
PS Claim 4; Page 341-343; 351pp; English.
XX
XX This sequence encodes a human protein of the invention, which has
CC hydrophobic domains. The DNA sequences can be used as a probe or as a
CC genetic marker. The protein can also be used as a marker, and to identify
CC potential genetic disorders. The DNA and protein can also be used as
CC nutritional sources or supplements. The protein exhibits cytokine, cell
CC proliferation, cell differentiation activities and induces production of
CC other cytokines in certain cell populations. The protein also exhibits
CC immune stimulating or immune suppressing activity. It can be used in the
CC treatment of various immune deficiencies and disorders, and to treat
CC infectious diseases caused by viral, bacterial, fungal or other
CC infections. The protein is also used for treating autoimmune disorders
CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid
CC arthritis. It is also useful in the treatment of allergic reactions and
CC conditions such as asthma, and in immune suppression after organ
CC transplantation. The protein is useful in regulation of haematopoiesis
CC and consequently in the treatment of myeloid or lymphoid cell
CC deficiencies. It is also used in compositions for tissue growth or
CC regeneration. The protein is also used in the treatment of osteoporosis
CC or osteoarthritis and in the treatment of periodontal disease and other
CC tooth repair processes. The protein is used in the treatment of nervous
CC system disorders such as Alzheimer's disease, Parkinson's disease, and
CC Huntington's disease. They are useful for protection or regeneration and
CC treatment of lung or liver fibrosis, reperfusion injury in various
CC tissues, and conditions resulting from systemic cytokine damage. They are
CC also used for promoting or inhibiting tissue differentiation. They are
CC also used as contraceptives since they exhibit activin or inhibin related
CC activities and as a fertility inducing therapeutic. They are used for
CC treating various coagulation disorders and in treatment and prevention of
CC conditions resulting from coagulation activities e.g. myocardial
CC infarction or stroke. They also act as receptors, receptor ligands or
CC inhibitors or agonists of receptor/ligand interactions. They are used to
CC treat inflammatory conditions such as septic shock, sepsis, ischaemia
CC reperfusion injury, arthritis, and nephritis. They can be used to
CC prevent tumours.
XX
XX Sequence 1527 BP; 382 A; 373 C; 361 G; 411 T; 0 other;
SQ
Query Match 97.7%; Score 1520; DB 21; Length 1527;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 35 GCAGTGGCCGTTACGGCGGAAAGATGGCGTCTTGGCACTCTAATTCGCTCGTGAT 94
DB 1 GCAGTGGCCGTTACGGCGGAAAGATGGCGTCTTGGCACTCTAATTCGCTCGTGAT 60
QY 95 TCGGTGGCCGACCTTTCACGATGGCTGCGCCCAACCTTACTACCTCTCTGCGCCCTGCTC 154
DB 61 TCGGTGGCCGACCTTTCACGATGGCTGCGCCCAACCTTACTACCTCTCTGCGCCCTGCTC 120
QY 155 TCTGTGCTCTCTCTACGTCGTCGAGAAACTGCGCGCGCTCTGCCAGGCTCTGCCACCCAA 214
DB 121 TCTGTGCTCTCTCTACGTCGTCGAGAAACTGCGCGCGCTCTGCCAGGCTCTGCCACCCAA 180
QY 215 CGGGAAGACGGTACCGCTGTGACTTTGACTGGAGAGAGTGGAGATCCTGATGTTCTC 274
DB

DB 181 CGCGAAGACGGTAACCCGTTGACITTTGACTGGAGAGAAGTGGAGATCCTGATGTTTCTC 240
QY 275 AGTGCCATTGTGATGATGAAGAACCGCAGATCCATGTTCTCTGATGACGTGCAAAACCCCC 334
DB 241 AGTGCCATTGTGATGATGAAGAACCGCAGATCCATGTTCTCTGATGACGTGCAAAACCCCC 300
QY 335 CTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAAACCACTTATGATGAGGAATA 394
DB 301 CTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAAACCACTTATGATGAGGAATA 360
QY 395 GAAACGGGACAAGAGGCTCACTTGGATTGTGGAGTCTTTGGCCAAATGGTCTTAATGACTGC 454
DB 361 GAAACGGGACAAGAGGCTCACTTGGATTGTGGAGTCTTTGGCCAAATGGTCTTAATGACTGC 420
QY 455 CAATCATTTGGCCCTATCATCTGACCTCTCCCTTAAATACAACTGTACAGGCTAAAT 514
DB 421 CAATCATTTGGCCCTATCTATCTGACCTCTCCCTTAAATACAACTGTACAGGCTAAAT 480
QY 515 TTTGGGAAGGTGATGTTGGACGCTATCTACTGATGTAGTACGGGTACAAAGTGACGACA 574
DB 481 TTTGGGAAGGTGATGTTGGACGCTATCTACTGATGTAGTACGGGTACAAAGTGACGACA 540
QY 575 TCACCCCTCACCAAGCAACTCCCTACCTGATCCTTCCAAAGGTGGCAAGGAGGCAATG 634
DB 541 TCACCCCTCACCAAGCAACTCCCTACCTGATCCTTCCAAAGGTGGCAAGGAGGCAATG 600
QY 635 CGCGGGCCACAGATTGACAAGAAAGGCGGCTGTCTCATGGACCTTCTCTGAGGAGAAT 694
DB 601 CGCGGGCCACAGATTGACAAGAAAGGCGGCTGTCTCATGGACCTTCTCTGAGGAGAAT 660
QY 695 GTGATCCGAGAAATTTAACTTAATGAGCTATACAGCGGGCCCAAGAACTATCAAAAGGT 754
DB 661 GTGATCCGAGAAATTTAACTTAATGAGCTATACAGCGGGCCCAAGAACTATCAAAAGGT 720
QY 755 GGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACCCACACGCTGTCAGATGG 814
DB 721 GGAGACAATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACCCACACGCTGTCAGATGG 780
QY 815 GAAACACAAGAGGATAAATAAGATCCTCACTTTGGCAGTCTTCTCTCTCTCAATTC 874
DB 781 GAAACACAAGAGGATAAATAAGATCCTCACTTTGGCAGTCTTCTCTCTCTCAATTC 840
QY 875 AGGCTCTTCCATAACCAACACAGCTGAGGCTGAGGCTGAGGCTTCTCTCTCTCTCTCTTGG 934
DB 841 AGGCTCTTCCATAACCAACACAGCTGAGGCTGAGGCTGAGGCTTCTCTCTCTCTCTCTTGG 900
QY 935 CTGTGACTGGTGGCGCAGCATGTCAGTCTCTGATTTTAAAGAGGCACTTAGGGAATGTC 994
DB 901 CTGTGACTGGTGGCGAGCATGTCAGTCTCTGATTTTAAAGAGGCACTTAGGGAATGTC 960
QY 995 AGGCACCTACAGGAAGGCTGCCATGCTGTGGCCAACCTGTTTCACTGGAGCAAGAAAGA 1054
DB 961 AGGCACCTACAGGAAGGCTGCCATGCTGTGGCCAACCTGTTTCACTGGAGCAAGAAAGA 1020
QY 1055 GATCTCATAGACGAGGAGGGAATGTTTCCCTCCAAGCTTGGGTGAGTGTGTTAACTG 1114
DB 1021 GATCTCATAGGAGGAGGGAATGTTTCCCTCCAAGCTTGGGTGAGTGTGTTAACTG 1080
QY 1115 CTTATCAGCTATTCAGACATCTCCATGCTTCTCCATGAAACTCTGTGGTTTCATCATTC 1174
DB 1081 CTTATCAGCTATTCAGACATCTCCATGCTTCTCCATGAAACTCTGTGGTTTCATCATTC 1140
QY 1175 CTTCTTACTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCTTAAGGTAGATGCTG 1234
DB 1141 CTTCTTACTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCTTAAGGTAGATGCTG 1200
QY 1235 GGGTATGAACGCTAAGAAATTTCCCAAGGACTCTTGGTTCCTTAAAGCCCTTCTGGCT 1294
DB 1201 GGGTATGAACGCTAAGAAATTTCCCAAGGACTCTTGGTTCCTTAAAGCCCTTCTGGCT 1260
QY 1295 TCGTTTATGTCCTTCAATTAAGATATAGCCTAACTTTGTCGCTAGTCTTAAAGGAAAC 1354
DB 1261 TCGTTTATGTCCTTCAATTAAGATATAGCCTAACTTTGTCGCTAGTCTTAAAGGAAAC 1320

QY 1355 CTTTAAACACAAAGTTTATCATTTAGACAAATATTGAACACCCCTATTGTTGGG 1414
Db CTTTAAACACAAAGTTTATCATTTAGACAAATATTGAACACCCCTATTGTTGGG 1380
QY 1415 ATTGAGAGGGTGAATAGAGGCTTGAGACTTTCTTTGTTGTTAGACTTTGGAGGAGA 1474
Db ATTGAGAGGGTGAATAGAGGCTTGAGACTTTCTTTGTTGTTAGACTTTGGAGGAGA 1440
QY 1475 AATCCCTCGACTTTTCACTAACCTCTGACATCTCCCAACACCCAGTTGATGCTTCC 1534
Db AATCCCTCGACTTTTCACTAACCTCTGACATCTCCCAACACCCAGTTGATGCTTCC 1500
QY 1535 GTAATAAAAGATTGGGATT 1554
Db GTAATAAAAGATTGGGATT 1520

RESULT 2
ID AAF16316
AAFI6316 standard; cDNA; 1598 BP.
AAFI6316;
13-MAR-2001 (first entry)
Human prostate cancer antigen nucleotide sequence SEQ ID NO:751.
Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
vulnerable; gastrointestinal; nephrotropic; antiinfective; gynaecological;
antibacterial; gene therapy; neural; immune; reproductive; renal;
gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
wound; infectious disease; ss.
Homo sapiens.
WO200055174-A1.
21-SEP-2000.
08-MAR-2000; 2000WO-US05988.
12-MAR-1999; 99US-0124270.
(HUMA-) HUMAN GENOME SCI INC.
(ROSE/) ROSEN C A.
Rosen CA, Ruben SM;
WPI; 2000-587513/55.
P-PSDB; AAB57113.
Prostate cancer associated gene sequences, referred to as prostate
cancer antigens, useful for treatment, prevention, and diagnosis of
disorders such as prostate cancer.
Claim 1; Page 1172; 2338pp; English.
AAFI5566 to AAFI6505 encode the human prostate cancer associated
proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
The prostate cancer antigens can have neuroprotective, cytostatic,
cardioactive, immunomodulatory, muscular, vulnerable, gastrointestinal,
nephrotropic, antiinfective, gynaecological and antibacterial activities,
and can be used in gene therapy. The prostate cancer antigen
polynucleotides may be used for detection of prostate cancer, chromosome
identification, as chromosome markers, and for numerous other diagnostic
or research purposes. The prostate cancer antigens may be used to treat
disorders such as neural, immune, muscular, reproductive,
gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to
AAB57303 represent sequences used in the exemplification of the present
invention.

XX
SQ Sequence 1598 BP; 437 A; 381 C; 365 G; 414 T; 1 other;
Query Match 97.2%; Score 1511; DB 21; Length 1598;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 1511; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 42 CCGTTACGGCCGAAAGATGCGGTCTTGGCACCTCTAATTTGCTCTCGTGTATTCGGTGC 101
Db CCGTTACGGCCGAAAGATGCGGTCTTGGCACCTCTAATTTGCTCTCGTGTATTCGGTGC 85
QY 102 CGGACCTTTCACGATGCTGCCCAACCTTACTACTCTCTCGGCCCTCTCTCTCTGCTG 161
Db CGGACCTTTCACGATGCTGCCCAACCTTACTACTCTCTCGGCCCTCTCTCTCTGCTG 145
QY 162 CTTTCTACTCTGAGGAACTGCCGCGCTCTGCCACGGTCTGCCACCGGAG 221
Db CTTTCTACTCTGAGGAACTGCCGCGCTCTGCCACGGTCTGCCACCGGAG 205
QY 222 ACGCAACCCGCTGTGACTTTGACTGGAGAGAGTGGAGATCTGTATTTCTCAGTGCA 281
Db ACGCAACCCGCTGTGACTTTGACTGGAGAGAGTGGAGATCTGTATTTCTCAGTGCA 265
QY 282 TTGTGATGAGAAAGACCGCAGATCCATGTTCTGATGACGTGCAAAACCCCTATATA 341
Db TTGTGATGAGAAAGACCGCAGATCCATGTTCTGATGACGTGCAAAACCCCTATATA 325
QY 342 TGGCCCTCTGATTTCAAGTACTTCAATGATAAAACCATTTGATGAGGAATAGAACGG 401
Db TGGCCCTCTGATTTCAAGTACTTCAATGATAAAACCATTTGATGAGGAATAGAACGG 385
QY 402 ACAAGAGGTCACCTGGATGCTGGAGTCTTTGGCAATTTGCTTAATGACGTGCCAATCAT 461
Db ACAAGAGGTCACCTGGATGCTGGAGTCTTTGGCAATTTGCTTAATGACGTGCCAATCAT 445
QY 462 TTGCCCTCTATCTATGCTGACCTTCCCTTAATCAACTGTACAGGCTTAATTTGGGA 521
Db TTGCCCTCTATCTATGCTGACCTTCCCTTAATCAACTGTACAGGCTTAATTTGGGA 505
QY 522 AGTGGATGTTGGACGCTATACTGATGTAGTACGGGTACAAAGTACAGCATCACCCC 581
Db AGTGGATGTTGGACGCTATACTGATGTAGTACGGGTACAAAGTACAGCATCACCCC 565
QY 582 TCACCAAGCACTCCCTACCTGATCTCTTCCAGGTGGCAAGGAGCAATGCGCGGC 641
Db TCACCAAGCACTCCCTACCTGATCTCTTCCAGGTGGCAAGGAGCAATGCGCGGC 625
QY 642 CACAGATTGACAAAGAGGCGGCTGTCTCATGACCTTCTCTGAGGAGATGTGATCC 701
Db CACAGATTGACAAAGAGGCGGCTGTCTCATGACCTTCTCTGAGGAGATGTGATCC 685
QY 702 GAGAAATTAATTAATGAGCTATACCAAGCGGCGCAAGAACTATCAAGGCTGGAGACA 761
Db GAGAAATTAATTAATGAGCTATACCAAGCGGCGCAAGAACTATCAAGGCTGGAGACA 745
QY 762 ATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACACAGTGTCAAGTGGGAAAACA 821
Db ATATCCCTGAGGAGCAGCTGTGGCTTCAACCCCAACACAGTGTCAAGTGGGAAAACA 805
QY 822 AAGAGGATAAATAGATCCTCATTGCGAGTGTCTCTCTCTGTCATTTCCAGGCTCT 881
Db AAGAGGATAAATAGATCCTCATTGCGAGTGTCTCTCTCTGTCATTTCCAGGCTCT 865
QY 882 TTCATTAACACAGCCTGAGGCTGACCTTTTATTTATGCTTTTCCCTTTGGCTGAC 941
Db TTCATTAACACAGCCTGAGGCTGACCTTTTATTTATGCTTTTCCCTTTGGCTGAC 925
QY 942 TGGGTGGGCGAGCATGACGCTTCTGATTTTAAAGAGGCATCTAGGGAATTTCTAGGCAAC 1001
Db TGGGTGGGCGAGCATGACGCTTCTGATTTTAAAGAGGCATCTAGGGAATTTCTAGGCAAC 985
QY 1002 CTACAGGAAGCCTGCCATCTGTGGCAACTGTTTCACTGGAGCAAGAGATCTCA 1061
Db CTACAGGAAGCCTGCCATCTGTGGCAACTGTTTCACTGGAGCAAGAGATCTCA 1045

QY 141 AspLeuSerLeuLysTyrAsnCysThrGlyLeuAsnPhcGlyLysValaspValGlyArg 160
|||||
Db 463 GACCTCTCCCTTAAATACAACTGTACAGGCTAAATTTGGGAAGGTGGATGTTGGACGC 522
QY 161 TyrThrAspValSerThrArgTyrLysValSerThrSerProLeuThrLysGlnLeuPro 180
|||||
Db 523 TATACCTGATGTAGTACGGGTACAAGAGGACATCATCCCTCACCAGCAACATCCCT 582
QY 181 ThrLeuLeuPheGlnGlyLysGluAlaMetArgProGlnLeuAspLysLys 200
|||||
Db 583 ACCCTGATCTCTTCAAGGTGCAAGGAGCAATCGCGGCCACAGATTGACAAGAA 642
QY 201 GlyArgAlaValSerThrPheSerGluGluAsnValleArgGluPheAsnLeuAsn 220
|||||
Db 643 GGACGGCTGTCTCATGGACCTTCTCTGAGGAGATGTGATCGGAGAAATTTAACTTAAAT 702
QY 221 GluLeuTyrGlnArgAlaLysLysLeuSerLysAlaGlyAspAsnLeuProGluGln 240
|||||
QY 703 GAGCTATACACGGGCGCCCAAGAACTATCAAGGCTGGAGACAATATCCCTGAGGAGCAG 762
QY 241 ProValAlaSerThrProThrThrValSerAspGlyGluAsnLysLysAspLys 258
|||||
Db 763 CCTGTGGCTTCAACCCCAACACAGTGTCTGAGTGGGGAACAAAGAGGTAA 816

RESULT 3

US-09-905-291A-206
Sequence 206, Application US/09905291A
Patent No. US2002016037A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kjaevan, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Paul, James
APPLICANT: Paoli, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: cids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905,291A
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-01-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 206
LENGTH: 1620
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (973)
OTHER INFORMATION: a, t, c or g
NAME/KEY: modified_base
LOCATION: (977)
OTHER INFORMATION: a, t, c or g
NAME/KEY: modified_base
LOCATION: (996)
OTHER INFORMATION: a, t, c or g
NAME/KEY: modified_base
LOCATION: (1003)
OTHER INFORMATION: a, t, c or g
US-09-905-291A-206

Alignment Scores:

Pred. No.: 1.84e-175 Length: 1620
Score: 1326.00 Matches: 257
Percent Similarity: 87.16% Conservative: 1
Best Local Similarity: 86.82% Mismatches: 0
Query Match: 97.57% Indels: 38
DB: 9 Gaps: 1

US-09-954-846-2 (1-258) x US-09-905-291A-206 (1-1620)

QY 1 MetAlaValLeuAlaProLeuLeuAlaLeuValTySerValProArgLeuSerArgTrp 20
|||||
Db 3 ATGGCGGTCTGGCACCTCTAAATGCTCTCGTGTATTGGTGGCGGCGACTTTCAGATGG 62
QY 21 LeuAlaGlnProTyrTyrLeuLeuSerAlaLeuLeuSerAlaAlaPheLeuValArg 40
|||||
Db 63 CTGCCCCAACCTTACTACCTTCTGCGGCCCTGCTCTCTGCGCTTCTCTACTCTGAGG 122
QY 41 LysLeuProLeuCysHisGlyLeuProThrGlnArgGluAspGlyAsnProCysasp 60
|||||
Db 123 AAACCTGCGCGGCTCTGCCACGGTCTGCCACCCACCCAGCGAAGACCGTAACCCGCTGTGAC 182
QY 61 PheAspTrpArgGluValGluLeuLeuMetPheLeuSerAlaAlaLeuValMetLysAsn 80
|||||
Db 183 TTTGACTGGAGAAAGTGGAGATCCTGATGTTTCTCAGTCCCATTTGATGATGATGAAGAAC 242
QY 81 ArgArgSer 83
|||||
Db 243 CGCAGATCCATCTACTGTGGAGACACATATAGGCAACATTTTCATGTTTAGTAAGTGGCC 302
QY 83 83
Db 303 AACACAATTCTTTCTTCGCTTGGATATTTCGCATGGGCGCTACTTTACATCACACTGCG 362

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT19
CLONE: 3244141
SEQUENCE DESCRIPTION: SEQ ID NO: 4 :
US-09-954-846-4

Query Match
Best Local Similarity 100.0%; Score 1555; DB 10; Length 1555;
Matches 1555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAGGCGGGGCGAGACCTACGACCGCGGCGAGCAGTGGCCGTACGGCGCGAAAAGAT 60
DB 1 AGGAGGCGGGGCGAGACCTACGACCGCGGCGAGCAGTGGCCGTACGGCGCGAAAAGAT 60
61 GGGGCTTTGGGCACTTAATTTGCTCTCGTGTATTCGGTGCGCGGACCTTTCACGATGGCT 120
DB 61 GGGGCTTTGGGCACTTAATTTGCTCTCGTGTATTCGGTGCGCGGACCTTTCACGATGGCT 120
121 CGCCCAACCTTACTACCTTCTGCGGCCCTGCTCTCTGCTGCTTCTACTCTCTGAGGAA 180
DB 121 CGCCCAACCTTACTACCTTCTGCGGCCCTGCTCTCTGCTGCTTCTACTCTCTGAGGAA 180
181 ACTGCGCGCGCTCTGCGGCCCTGCTCTCTGCTGCTTCTACTCTCTGAGGAA 240
DB 181 ACTGCGCGCGCTCTGCGGCCCTGCTCTCTGCTGCTTCTACTCTCTGAGGAA 240
241 TGACTGGAGAGAGTGGAATCTGATGTTTCTCAGTGCCCATTTGATGATGAAGAACC 300
DB 241 TGACTGGAGAGAGTGGAATCTGATGTTTCTCAGTGCCCATTTGATGATGAAGAACC 300
301 CAGATCCATCTTCTGATGAGTGAACCAACCCCTTATATATGCGCCCTGAGTATCA 360
DB 301 CAGATCCATCTTCTGATGAGTGAACCAACCCCTTATATATGCGCCCTGAGTATCA 360
361 GTACTTCAATGATAAACCATTGATGAGGAGTGAACGAGGAGGTCACCTTGAT 420
DB 361 GTACTTCAATGATAAACCATTGATGAGGAGTGAACGAGGAGGTCACCTTGAT 420
421 TGTGGAGTCTTTGGCAATTTGCTTAATGATGCGCAATCATTTGCCCTTACTATGCTGA 480
DB 421 TGTGGAGTCTTTGGCAATTTGCTTAATGATGCGCAATCATTTGCCCTTACTATGCTGA 480
481 CCTCTCCCTTAATACAACTGACAGGCTAAATTTGGGAAGTGTGTCAGCCTA 540
DB 481 CCTCTCCCTTAATACAACTGACAGGCTAAATTTGGGAAGTGTGTCAGCCTA 540
541 TACTGATGTTAGTACGCGGTACAAAGTGAGCAGATCACCTACCAAGCAACTCCCTAC 600
DB 541 TACTGATGTTAGTACGCGGTACAAAGTGAGCAGATCACCTACCAAGCAACTCCCTAC 600
601 CTTGATCTCTTCAAGGTGGCAAGGAGGCAATGCGCGGCGCAAGATTGACAAAGG 660
DB 601 CTTGATCTCTTCAAGGTGGCAAGGAGGCAATGCGCGGCGCAAGATTGACAAAGG 660
661 ACGGCTGCTCTATGACCTTCTCTGAGGAGATGATCGGAGATTTACITTAATGA 720
DB 661 ACGGCTGCTCTATGACCTTCTCTGAGGAGATGATCGGAGATTTACITTAATGA 720
721 GCTATACCAAGCGGCGCAAGAACTATCAAGGCTGGAGCAATATCCCTGAGGAGCACC 780
DB 721 GCTATACCAAGCGGCGCAAGAACTATCAAGGCTGGAGCAATATCCCTGAGGAGCACC 780
781 TGTGGCTTCAACCCCAACCACTGTCAGATGGGGAACCAAGAGGATAAATGAATCC 840
DB 781 TGTGGCTTCAACCCCAACCACTGTCAGATGGGGAACCAAGAGGATAAATGAATCC 840
841 TCACCTTTGGCAGTGTCTCTCTCTGTCATATTCAGGCTTTTCATACCAAGCCTG 900
DB 841 TCACCTTTGGCAGTGTCTCTCTCTGTCATATTCAGGCTTTTCATACCAAGCCTG 900

RESULT 2

US-09-925-300-751
; Sequence 751, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 751
; LENGTH: 1598
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-300-751

Query Match 97.2%; Score 1511; DB 10; Length 1598;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1511; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 42 CCCTAGCGGCGGCAAGAGATGCGGCTTTCGACCTCTAATGCTCTCGTATTCGGTGC 101

Db 26 CGGTACGGCCGAAAGATGGGGTCTTGGACCTCTAATGCTCTCGTGTATTCGGTGC 85
Qy 102 CGGACTTTCACGATGGCTCGCCCAACCTTACTACTCTCTGCGGCCCTGCTCTCTGCTG 161
Db 86 CGGACTTTCACGATGGCTCGCCCAACCTTACTACTCTCTGCGGCCCTGCTCTCTGCTG 145
Qy 162 CTTCTCTACTCGTGAAGAAATGCGCGCGCTCTGCGACGGTCTGCCCAACCCCAACGCGAAG 221
Db 146 CTTCTCTACTCGTGAAGAAATGCGCGCGCTCTGCGACGGTCTGCCCAACCCCAACGCGAAG 205
Qy 222 ACCTTAACCCCTGTGACTTTGACTGGAGAGAGTGGAGATCTGATGTTCTCAGTGCCA 281
Db 206 ACCTTAACCCCTGTGACTTTGACTGGAGAGAGTGGAGATCTGATGTTCTCAGTGCCA 265
Qy 282 TTGTGATGATGAAGAACCGCAGATCCATGTTCTGATGACGTGCAAAACCCCTCTATATA 341
Db 266 TTGTGATGATGAAGAACCGCAGATCCATGTTCTGATGACGTGCAAAACCCCTCTATATA 325
Qy 342 TGGCCCTGAGTATATCAAGTACTTCAATGATATAAACCATTGATGAGGAACCTAGAAGGG 401
Db 326 TGGCCCTGAGTATATCAAGTACTTCAATGATATAAACCATTGATGAGGAACCTAGAAGGG 385
Qy 402 ACAAGAGGCTCAGTTGGATTTGGAGTCTTTGCCAATTTGGTCTTAATGACTGCCAATCAT 461
Db 386 ACAAGAGGCTCAGTTGGATTTGGAGTCTTTGCCAATTTGGTCTTAATGACTGCCAATCAT 445
Qy 462 TTGCCCTTATCTATGCTGACCTCTCCCTTAATACAACTGTACAGGGCTAAATTTGGGA 521
Db 446 TTGCCCTTATCTATGCTGACCTCTCCCTTAATACAACTGTACAGGGCTAAATTTGGGA 505
Qy 522 AGGTGGATGTGGAGCGCTATCTATGTTAGTACGGGTACAAAGTGAGCACATCACCC 581
Db 506 AGGTGGATGTGGAGCGCTATCTATGTTAGTACGGGTACAAAGTGAGCACATCACCC 565
Qy 582 TCACCAAGCAACTCCCTACCTCTGATGTCCTTCCAAAGTGGCAAGGAGGCAATGCGCGGC 641
Db 566 TCACCAAGCAACTCCCTACCTCTGATGTCCTTCCAAAGTGGCAAGGAGGCAATGCGCGGC 625
Qy 642 CACAGATGACAAGAAAGAGGGCTGCTCATGACCTTCTCTGAGGAGATGTATCC 701
Db 626 CACAGATGACAAGAAAGAGGGCTGCTCATGACCTTCTCTGAGGAGATGTATCC 685
Qy 702 GAGATTTAACTTAATGAGCTATACACGCGGCCCAAGAAACTCAAAAGGCTGGAGACA 761
Db 686 GAGATTTAACTTAATGAGCTATACACGCGGCCCAAGAAACTCAAAAGGCTGGAGACA 745
Qy 762 ATATCCCTGAGGAGCAGCTGTGGGTTCAACCCCAACCCACACAGTGTAGATGGGAAACA 821
Db 746 ATATCCCTGAGGAGCAGCTGTGGGTTCAACCCCAACCCACACAGTGTAGATGGGAAACA 805
Qy 822 AGAAGGATAAATAGATCCTCATTGGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 881
Db 806 AGAAGGATAAATAGATCCTCATTGGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 865
Qy 882 TTCCATAACCAAGCCTGAGGCTGCAGCTTTTATTTATGTTTCCCTTTGGCTGTGAC 941
Db 866 TTCCATAACCAAGCCTGAGGCTGCAGCTTTTATTTATGTTTCCCTTTGGCTGTGAC 925
Qy 942 TGGGTGGGGCAGCATGCAAGCTTCTGATTTTAAAGAGGCAATGAGGAATTTGTCAGGCACC 1001
Db 926 TGGGTGGGGCAGCATGCAAGCTTCTGATTTTAAAGAGGCAATGAGGAATTTGTCAGGCACC 985
Qy 1002 CTACAGGAGGCTGCGCATGCTGTGGCCCAACTGTTTCACTGGAGCAAGAAAGAGATCTCA 1061
Db 986 CTACAGGAGGCTGCGCATGCTGTGGCCCAACTGTTTCACTGGAGCAAGAAAGAGATCTCA 1045
Qy 1062 TAGGACGAGGGGAAATGGTTTCCCTCAAGCTTGGGTGAGTGTAACTGCTATCA 1121
Db 1046 TAGGACGAGGGGAAATGGTTTCCCTCAAGCTTGGGTGAGTGTAACTGCTATCA 1105
Qy 1122 GCTATTGAGAGATCTCCATGCTTTTCCATGAACTCTGTGTTTCAATCATCTCTCTTA 1181

Db 1106 GCTATTTCAGACATCTCCATGGTTTCTCCATGAACACTCTCTGTTTTCATCATCTCTCTTA 1165
Qy 1182 GTTGACCTTCACAGCTTGTGTAGACCTAGATTTAAACCTTAAGTAAGATGCTGGGTATA 1241
Db 1166 GTTGACCTTCACAGCTTGTGTAGACCTAGATTTAAACCTTAAGTAAGATGCTGGGTATA 1225
Qy 1242 GAACGCTAAGAAATTTTCCCCCAAGGACTCTGCTTCCCTTAAGCCCTTCTGGCTTCGTTTA 1301
Db 1226 GAACGCTAAGAAATTTTCCCCCAAGGACTCTGCTTCCCTTAAGCCCTTCTGGCTTCGTTTA 1285
Qy 1302 TGGTCTTCAATTAAGTAAGAACTTAAGCTTAAGCTTCTGCTTCCCTTAAGCCCTTCTGGCTTCGTTTA 1361
Db 1286 TGGTCTTCAATTAAGTAAGAACTTAAGCTTCTGCTTCCCTTAAGCCCTTCTGGCTTCGTTTA 1345
Qy 1362 CACAAGATTTTATCATTTGAAGACATATTTGAACAACCCCTATTTTGTGGGATTCGAGA 1421
Db 1346 CACAAGATTTTATCATTTGAAGACATATTTGAACAACCCCTATTTTGTGGGATTCGAGA 1405
Qy 1422 AGGGTGAATAGAGGCTTGAGACTTTTCTTGTGTGTAGGACTTGGAGGAGAAATCCCC 1481
Db 1406 AGGGTGAATAGAGGCTTGAGACTTTTCTTGTGTGTAGGACTTGGAGGAGAAATCCCC 1465
Qy 1482 TGGACTTTTCACTAACCTCTGACATACTCCCAACCCCTGATGATGGCTTCCGTAATAA 1541
Db 1466 TGGACTTTTCACTAACCTCTGACATACTCCCAACCCCTGATGATGGCTTCCGTAATAA 1525
Qy 1542 AAAGATTTGGGATT 1554
Db 1526 AAAGATTTGGGATT 1538

RESULT 3
US-09-852-797-19
Sequence 19, Application US/09852797
Patent No. US20020172994A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 28 Human Secreted Proteins
FILE REFERENCE: P2003P2
CURRENT APPLICATION NUMBER: US/09/852,797
CURRENT FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: 60/265,583
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 09/152,060
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: PCT/US98/04858
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/040,762
PRIOR FILING DATE: 1997-03-14
PRIOR APPLICATION NUMBER: 60/040,710
PRIOR FILING DATE: 1997-03-14
PRIOR APPLICATION NUMBER: 60/050,934
PRIOR FILING DATE: 1997-03-30
PRIOR APPLICATION NUMBER: 60/048,100
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: 60/048,357
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: 60/048,189
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: 60/057,765
PRIOR FILING DATE: 1997-09-05
PRIOR APPLICATION NUMBER: 60/048,970
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/068,368
PRIOR FILING DATE: 1997-12-19
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 1699
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE

